

UNIVERSITE JOSEPH FOURIER-GRENOBLE 1

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ JOSEPH FOURIER

Discipline : Biologie Structurale

Présentée et soutenue publiquement par

Clothilde MANZANO

Le 07 Septembre 2009

**Caractérisation structurale et fonctionnelle des
composants du pilus de *Streptococcus pneumoniae* : vers
une meilleure compréhension de la biogenèse des pili.**

Dr. Anne IMBERTY

Présidente de jury

Prof. Laurent GUTMANN

Rapporteur

Dr. Vincent VILLERET

Rapporteur

Dr. Patricia RENESTO

Invité

Dr. Andréa DESSEN

Directrice de thèse

Thèse préparée au Laboratoire des Protéines Membranaires (LPM) de l'Institut de Biologie
Structurale Jean-Pierre EBEL (IBS) CEA/CNRS/UJF

Remerciements

Je tiens tout d'abord à remercier Andréa Dessen pour m'avoir accueillie dans son laboratoire pendant ces 3 ans. Merci de la confiance que tu m'as accordée, de m'avoir fait bifurquer de sujet de thèse et ainsi de m'avoir fait découvrir le monde passionnant des pili pneumococciques!

Un merci particulier à petit biquet ! Thierry, merci, d'avoir été mon souffre-douleur ! Je suis heureuse de t'avoir rencontrée ; cela aurait été beaucoup moins drôle si tu n'avais pas rejoint l'équipe. Les moments de discussion, scientifiques ou non d'ailleurs, vont me manquer... Et continue à bricoler, comme ça je ferais appel à toi pour mes futurs travaux, mais attention, pas de dédommagements si la malédiction du pouce te poursuit !

Merci à mon autre compère de thèse, PJ, et courage pour la fin de ta thèse et ton ouverture de boulangerie, je serais première à venir commander de bons gâteaux ! Merci pour les moments de franches rigolades ! Merci Vivi pour tes conseils, ta disponibilité et ta gentillesse. Les moments détentes hammam + massages vont me manquer ! Bref sans vous 3, Titi, PJ et Vivi, la vie au labo aurait été bien plus triste...

Je remercie Patricia Renesto, Anne Imberty, Laurent Gutmann et Vincent Villeret pour avoir accepté de juger ce travail. Merci du temps que vous y avez consacré.

Merci à tous les membres du labo auprès de qui j'ai passé des années agréables. Manu, merci pour les réelles discussions d'avenir que nous avons eues. Merci Cécile, ta présence au labo m'a souvent manquée. Merci David, bonne chance pour ton retour au Brésil. Et merci aux autres membres du LPM...

Je tiens à remercier également les membres du LIM, venant souvent vous voir et emprunter du matériel ou vous demander conseil, vous ne m'avez jamais rejetée. Merci en particulier à Lamya, Cécile, Julien, Laure, je vous souhaite bonheur et réussite.

Merci à Jean-Pierre (Simorre), Thomas et Cécile pour l'étude RMN de la sortase. Même si ce ne fut que de préliminaires études, cela fut agréable de travailler et de discuter avec vous, les autres structuralistes ! Merci Guy (Schoehn) et Daphna pour votre gentillesse et votre disponibilité quant aux études de microscopie électronique

Merci à mes parents pour votre soutien pendant toutes ces (longues ?!) années d'études. Ma ptite Mamounette, je t'ai bien bassinée avec mes bactéries et mes protéines durant ces 3 années ! Heureusement que tu veux retourner sur les bancs de la fac d'histoire de l'art ou de langues, sinon pas besoin de postuler pour la fac de bio, tu es déjà au point ! Merci Charlou, même à distance, tu as été d'un grand soutien pour moi.

Et finalement, merci Bertrand, sans toi rien n'aurait été pareil et je pense que j'aurais eu bien du mal à finir cette thèse. J'aurais tellement de choses à te dire et à te remercier alors, simplement, merci pour ton amour.

Table des matières

ABREVIATIONS	1
INTRODUCTION.....	3
I. <i>Streptococcus pneumoniae</i>	4
I.1. Présentation.....	4
I.2. Un pathogène pour l'homme	5
I.2.1. Les infections associées	5
I.2.2. Le processus de l'infection	5
I.3. Les traitements disponibles.....	7
I.3.1. Les vaccins.....	7
I.3.2. Les antibiotiques.....	9
II. Les facteurs de virulence du pneumocoque.....	10
II.1. Les polysaccharides capsulaires	11
II.2. La paroi bactérienne.....	11
II.3. Les protéines pneumococciques.....	12
II.3.1. La protéine de surface PspA.....	12
II.3.2. L'adhésine CbpA	12
II.3.3. La pneumolysine	13
II.3.4. La neuraminidase	13
II.3.5. La hyaluronidase	13
II.3.6. L'autolysine LytA	14
II.4. Les pili.....	14
III. Les pili, organelles de surface bactériens	16
III.1. Les pili des bactéries à Gram-négatif	16
III.2. Les pili des bactéries à Gram-positif	18
III.2.1. Structure générale	18
III.2.2. Les sortases, enzymes catalysant la formation du pilus	19
III.2.3. Les pilines majeures et auxiliaires.....	20
III.2.4. Quelques exemples de pili de bactéries à Gram-positif	22
III.2.4.1. <i>Corynebacterium diphtheriae</i>	22
III.2.4.2. <i>Streptococcus agalactiae</i>	23
III.2.4.3. <i>Streptococcus pyogenes</i>	25
III.2.4.4. <i>Streptococcus pneumoniae</i>	26
IV. Les sortases et l'art d'accrocher des protéines cibles à la paroi bactérienne	30
IV.1. Classification des sortases.....	32
IV.1.1. Les sortases de classe A (SrtA)	32
IV.1.2. Les sortases de classe B (SrtB).....	34
IV.1.3. Les sortases de classe C (SrtC).....	35
IV.1.4. Les sortases de classe D (SrtD)	37
IV.2. Informations structurales sur les sortases ancrant des protéines cibles sur la paroi bactérienne.....	37
IV.2.1. Le repliement des sortases de classes A et B.....	37
IV.2.2. L'historique controversé du site actif.....	39
IV.2.3. La triade catalytique	42
CONTEXTE ET OBJECTIF DES TRAVAUX.....	44
MATERIEL ET METHODES	47
I. Biologie moléculaire	48
I.1. Clonage.....	48

I.2.	Mutagenèse dirigée.....	48
I.2.1.	Construction des mutants SrtC-1C193A, SrtC-1H131D, SrtC-R202E, SrtC-1D58GW60G, SrtC-1D58GW60GC193A.....	48
I.2.2.	Construction des mutants SrtC-1 _{lid2} et SrtC-2 _{lid1}	49
II.	Biochimie.....	50
II.1.	Surexpression de la protéine d'intérêt.....	50
II.2.	Purification de la protéine d'intérêt.....	50
II.2.1.	Chromatographie d'affinité sur colonne HisTrap TM HP.....	50
II.2.2.	Coupage de l'étiquette hexahistidine.....	51
II.2.3.	Chromatographie d'exclusion de taille.....	51
II.2.4.	Production de SrtC-1 et RrgB sélénées.....	51
II.3.	Formation <i>in vitro</i> des fibres de RrgB.....	52
III.	Analyses biophysiques.....	53
III.1.	Spectrométrie de masse.....	53
III.2.	Séquençage protéique N-terminal.....	54
III.3.	Étude de la thermostabilité des protéines par fluorescence : Thermal Shift Assay (TSA).....	54
III.3.1.	Principe.....	54
III.3.2.	Acquisition des données.....	55
III.4.	Microscopie électronique.....	56
III.4.1.	Principe.....	56
III.4.2.	Analyse des fibres de RrgB.....	56
IV.	Études cristallographiques des protéines SrtC-1, SrtC-3 et RrgB.....	58
IV.1.	Introduction à la cristallographie des rayons X.....	58
IV.1.1.	Cristallogénèse.....	58
IV.1.2.	Le principe de diffraction des rayons X par un cristal de protéine.....	59
IV.1.2.1.	Note préliminaire sur les cristaux de protéines.....	59
IV.1.2.2.	Diffraction des rayons X par les protéines du cristal.....	60
IV.1.2.3.	Facteur de structure et densité électronique.....	62
IV.1.3.	Détermination des intensités de diffraction.....	63
IV.1.4.	Résolution du problème de phase.....	64
IV.1.4.1.	Préambule : la loi de Friedel.....	64
IV.1.4.2.	La diffusion anormale.....	65
IV.1.4.3.	Le remplacement moléculaire.....	68
IV.1.5.	Construction du modèle et affinement.....	68
IV.1.6.	Analyse des structures.....	69
IV.2.	Cristallogénèse et cristallographie de SrtC-1.....	70
IV.2.1.	Cristallogénèse de SrtC-1.....	70
IV.2.2.	Enregistrement des données de diffraction.....	70
IV.2.3.	Structure de SrtC-1.....	70
IV.3.	Cristallogénèse et cristallographie de SrtC-3.....	71
IV.3.1.	Cristallogénèse de SrtC-3.....	71
IV.3.2.	Structure de SrtC-3.....	71
IV.4.	Cristallogénèse et cristallographie de RrgB.....	72
IV.4.1.	Cristallogénèse de RrgB.....	72
IV.4.2.	Enregistrement des données de diffraction.....	72
	RESULTATS ET DISCUSSION.....	73
I.	Etude fonctionnelle des composants du pilus de <i>S. pneumoniae</i>.....	74
I.1.	Rôle des sortases SrtC-1, SrtC-2 et SrtC-3 dans la polymérisation de RrgB.....	74
I.1.1.	Analyse <i>in vitro</i>	74
I.1.2.	Analyse <i>in vivo</i>	75
I.2.	Caractérisation des fibres de RrgB formées <i>in vitro</i>	77
I.2.1.	Purification des fibres.....	77
I.2.2.	Spectrométrie de masse des fibres.....	78
I.2.3.	Microscopie électronique des fibres.....	80

II. Détermination de la structure des sortases SrtC-1 et SrtC-3	83
II.1. Structure de SrtC-1	83
II.1.1. Production de SrtC-1	83
II.1.2. Cristallogenèse de SrtC-1	83
II.1.3. Production de cristaux de SrtC-1 sélénée	84
II.1.4. Résolution de la structure de SrtC-1	84
II.1.4.1. Collecte des données natives et sélénées	84
II.1.4.2. Traitement des données	85
II.1.5. Analyse de la structure de SrtC-1	86
II.1.5.1. Structure globale de SrtC-1	86
II.1.5.2. Site actif de SrtC-1	89
II.2. Structure de SrtC-3	91
II.2.1. Production de SrtC-3	91
II.2.2. Cristallogenèse	91
II.2.3. Résolution de la structure	91
II.2.4. Analyse de la structure de SrtC-3	94
II.2.4.1. Structure globale de SrtC-3	94
II.2.4.2. Site actif	95
II.3. La structure des sortases formant le pilus se distinguent des structures de sortases classiques	96
III. Caractérisation de la région du site actif de SrtC-1.....	100
III.1. Étude du couvercle de SrtC-1	100
III.1.1. Activité du mutant du couvercle SrtC-1D58GW60G	100
III.1.2. Stabilité thermique des mutants du couvercle SrtC-1D58GW60G et SrtC-1D58GW60G	101
III.1.3. Analyse des mutants SrtC-1_lid2 et SrtC-2_lid1	103
III.2. Étude du site actif de SrtC-1	104
III.2.1. Activité des mutants du site actif	104
III.2.2. Stabilité thermique des mutants	105
IV. Formation <i>in vitro</i> d'un complexe covalent entre SrtC-1 et RrgB.....	108
V. Étude structurale de RrgB.....	111
CONCLUSIONS ET PERSPECTIVES.....	114
REFERENCES BIBLIOGRAPHIQUES.....	120
ANNEXES	133
PUBLICATIONS	137

Abréviations

Abréviations

A	Absorbance
Å	Angström
ADN	Acide désoxyribonucléique
BET	Bromure d'éthidium
CAPS	Acide 3-cyclohexylamino-1-propane-sulfonique
CNRP	Centre national de référence des pneumocoques
CWSS	Cell wall sorting signal
DO	Densité optique
DTT	1,4-dithiothréitol
EDTA	Acide éthylène diamine tétraacétique
Ig	Immunoglobuline
IPTG	Isopropyl- β -D-thiogalactoside
LC MS/MS	Liquid chromatography coupled to tandem mass spectrometry
m/z	Rapport masse sur charge
MAD	Multiwavelength anomalous dispersion
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyéthylène glycol
PMSF	Phénylméthylsulfonyl fluoride
Qsp	Quantité suffisante pour
Rpm	Rotations par minute
Rmsd	Root mean square deviation
SAD	Single anomalous dispersion
SDS-PAGE	Gel d'électrophorèse en conditions dénaturantes
SIDA	Syndrome de l'ImmunoDéficiency Acquis
TSA	Thermal shift assay
U.A.	Unité d'absorbance

Introduction

I. *Streptococcus pneumoniae*

I.1. Présentation

En 1881, Sternberg, aux Etats-Unis, et Pasteur, en France, ont décrit indépendamment la présence d'une bactérie en forme de lancette, organisée en paire, dans des échantillons de salive. Alors qu'ils avaient identifié le même organisme, Sternberg le nomma *Micrococcus pasteurii* (Sternberg, 1881) et Pasteur, *Microbe septicémique de la salive* (Pasteur, 1881). Ce n'est qu'en 1974, après avoir été nommé *Pneumococcus*, en raison de sa capacité à causer des infections du système respiratoire, puis *Diplococcus pneumoniae*, que cette bactérie reçut le nom que nous lui connaissons aujourd'hui, *Streptococcus pneumoniae* (Deibel and Seeley, 1974).

S. pneumoniae, communément appelé pneumocoque, est une bactérie à Gram-positif, existant en cellule isolée, en paires ou sous forme de chaînettes (figure 1).

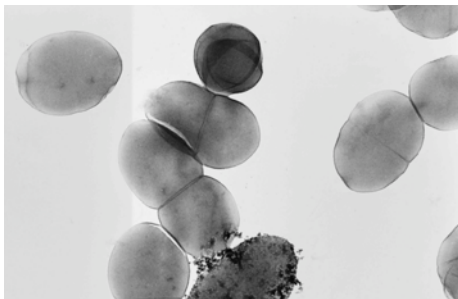


Figure 1 : Cliché de microscopie électronique de la souche *S. pneumoniae* non encapsulée R6 prise au Laboratoire de Microscopie Electronique de l' IBS.

Une des caractéristiques importantes du pneumocoque est qu'il peut être encapsulé, et dès lors, devenir pathogène. Cette capsule est composée de polysaccharides mais également, dans certains cas, d'acides, de ribitol ou d'arabinitol. La nature chimique de ces polysaccharides capsulaires peut varier. Ainsi, 91 variétés différentes de capsules sont connues aujourd'hui, ce qui a permis de caractériser autant de sérotypes (Henrichsen, 1995 ; Park *et al.*, 2007). Ces sérotypes n'ont pas tous le même pouvoir pathogène. La capsule polysaccharidique constitue un déterminant antigénique ; les sous-unités glucidiques la composant sont en effet reconnues par les anticorps humains pouvant ainsi conférer une protection immune.

I.2. Un pathogène pour l'homme

En 2005, l'Organisation Mondiale de la Santé a évalué à près de 1,6 million le nombre de personnes décédant chaque année d'une pneumococcie, parmi lesquelles 0,7 million d'enfants de moins de 5 ans, dont la plupart vivait dans des pays en voie de développement (<http://www.who.int/>).

I.2.1. Les infections associées

S. pneumoniae est présent à l'état commensal dans les voies respiratoires supérieures dès la prime enfance et est considéré comme un constituant normal de cette flore. Il colonise généralement le nasopharynx de façon asymptomatique. Toutefois, sous certaines conditions, le pneumocoque est responsable d'infections locales, sans gravité, mais aussi de pathologies sévères pouvant être mortelle. Deux types de maladies pneumococciques peuvent être distinguées : d'une part, les infections mucosales comprenant la majorité des maladies associées au pneumocoque, comme les otites, les sinusites, les bronchites ou les pneumonies, et, d'autre part, les maladies invasives. En effet, une partie des infections se limitant aux voies respiratoires peuvent progresser et causer de graves pathologies telles la méningite, la bactériémie ou la pneumonie nécrotique. Plusieurs facteurs de l'hôte sont associés à un risque accru de développer des maladies invasives. Ainsi, les infections se développent préférentiellement chez les enfants de moins de 2 ans et les personnes âgées de plus de 65 ans. De plus, les personnes immunodéprimées, atteintes d'un cancer ou du SIDA, sont également des populations à risque. En France, *S. pneumoniae* est la première cause de mortalité par infection bactérienne communautaire avant 2 ans avec 11,8% de mortalité en ce qui concerne les cas de méningite bactérienne et plus de 30% de séquelles lourdes associées (chiffres de l'Observatoire national des méningites bactériennes de l'enfant, de janvier 2001 à avril 2003).

I.2.2. Le processus de l'infection

S. pneumoniae pénètre chez l'hôte et colonise l'épithélium de la muqueuse du nasopharynx. Au sein des collectivités d'enfants, près de 50% des individus sont porteurs de cette bactérie et le taux de colonisation baisse sensiblement avec l'âge pour atteindre 5 à 10%

chez l'adulte (Levy, 1998). L'adhésion au niveau du nasopharynx est médiée par différents facteurs. Il a été ainsi montré l'importance des glycoconjugués présents à la surface des cellules pharyngeales mais aussi celle de la phosphorylcholine bactérienne pour l'adhérence du pneumocoque (Andersson *et al.*, 1983 ; Geelen *et al.*, 1993). Il a été également montré que les souches les plus adhérentes sont associées à des infections locales, telles la sinusite ou l'otite, alors que les souches moins adhérentes sont plutôt susceptibles de provoquer des bactériémies et méningites (Andersson *et al.*, 1981).

Les circonstances permettant ensuite au pneumocoque de migrer du nasopharynx vers les poumons restent encore mal comprises. Il se peut que l'état immunitaire de l'hôte au moment de la colonisation, tout autant que la virulence de la souche, déterminent si la bactérie demeurera au niveau du nasopharynx ou si elle envahira son hôte. Il a été néanmoins décrit que deux composés bactériens sécrétés, la pneumolysine et la protéase IgA1, facilitaient l'accès du pathogène aux poumons (Boulnois, 1992 ; Musher, 1992).

L'accès dans le sang est par la suite facilité par les dommages causés par la pneumolysine et le peroxyde d'hydrogène au niveau du poumon. Du sang, il peut ensuite migrer aux méninges et, sous l'action des neuraminidases et hyaluronidases dégradant le tissu conjonctif et permettant ainsi le passage de la barrière hémato-encéphalique, donner lieu à des méningites (Eberhardt *et al.*, 2003). La figure 2 ci-dessous résume le parcours infectieux du pneumocoque.

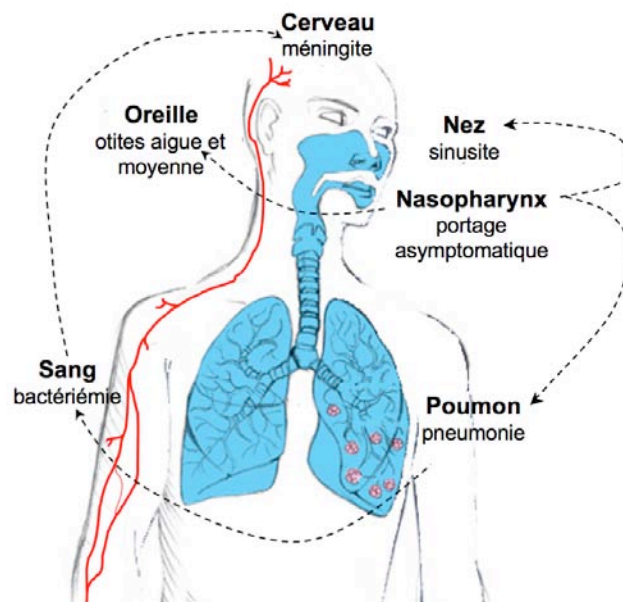


Figure 2 : Illustration schématique de l'infection pneumococcique dans l'hôte humain. La colonisation du pneumocoque au niveau du nasopharynx peut s'étendre jusqu'au nez, aux oreilles ou encore jusqu'aux poumons et provoquer respectivement des sinusites, des otites ou des pneumonies. Généralement, la bactériémie se passe tôt dans l'infection via les poumons alors que la méningite se passe après la bactériémie.

I.3. Les traitements disponibles

I.3.1. Les vaccins

La première tentative de vaccination, réalisée au début du 20^{ème} siècle et préparée à partir d'extrait bruts de pneumocoques, ne présenta qu'un rare effet protecteur et entraîna une forte réactivité responsable d'infections. Dans les années 1920, il fut montré que les polysaccharides de la capsule étaient antigéniques, ouvrant ainsi la voie des vaccins polysaccharidiques (Robbins *et al.*, 1983). Les premiers essais de vaccination furent alors effectués au début de la Seconde Guerre Mondiale sur des volontaires de l'armée américaine et aboutirent, en 1945, à un premier vaccin dirigé contre 4 polysaccharides (MacLeod, 1945). Toutefois, les premiers essais de vaccination furent assez rapidement suspendus en raison de l'apparition des antibiotiques. Ce ne fut qu'au milieu des années 1960 que les polysaccharides capsulaires retinrent de nouveau l'attention en tant que candidat pour un vaccin. En effet, malgré l'antibiothérapie, la gravité et la létalité de l'infection pneumococcique demeuraient préoccupantes en raison de l'émergence de la résistance aux antibiotiques. Une vaste campagne de vaccination fut entreprise avec succès en Afrique du Sud à l'aide de vaccins multivalents. Elle aboutit en 1977 à la création d'un nouveau vaccin à 14 valences (Austrian, 1981 ; Robbins *et al.*, 1983). Ce vaccin fut élargi à 23 valences en 1983 (Pneumovax®23-Merck, Pneumo®23-Aventis Pasteur et Pnu-Immune®-Wyeth Pharmaceuticals) couvrant environ 90% des infections causées par *S. pneumoniae* aux Etats-Unis (Robbins *et al.*, 1983). Malgré une efficacité certaine chez l'adulte, ce vaccin reste peu efficace chez l'enfant de moins de 2 ans, la personne âgée et la personne immunodéficiente (Douglas *et al.*, 1983) (Prevention of pneumococcal disease : recommendations of the Advisory Committee on Immunization Practices, 1997). Cette vaccination polysaccharidique ne provoque, en effet, qu'une faible réponse immunitaire chez l'enfant de bas âge. Seule une réponse B-lymphocytaire sans coopération cellulaire T est induite avec ce vaccin, or, l'enfant en bas âge possède des lymphocytes B immatures qui ne peuvent être activés et moduler une réponse immunitaire.

Une solution vaccinale alternative a alors été développée pour les enfants de moins de 2 ans. Il s'agit d'un vaccin, dit conjugué, qui associe les polysaccharides à une protéine de couplage, laquelle stimule plus efficacement les lymphocytes T, qui sont, eux, matures dès les

premières années de la vie. Le seul vaccin conjugué disponible sur le marché européen depuis 2001 (Prenar®-Wyeth Pharmaceuticals) contient les sérotypes des 7 souches les plus couramment répandues dans les infections invasives de l'enfant. Chaque polysaccharide de ce vaccin heptavalent est couplé à une protéine diphtérique non toxique. Les études menées aux Etats-Unis ont montré une diminution de 75 % de l'incidence des infections invasives chez les enfants de moins de 5 ans, 4 années après l'introduction du vaccin heptavalent (Centers for Disease Control and Prevention, 2005). En France, d'après les données du réseau EPIBAC, réseau de laboratoires de microbiologie permettant la surveillance des infections invasives communautaires notamment à pneumocoques, recueillies entre 2001 et 2006, la diminution était de 25 % pour les méningites et de 20 % pour les bactériémies chez les enfants de moins de 2 ans. Dans le même temps, l'incidence des méningites et bactériémies à pneumocoques ne diminuait pas chez les enfants plus âgés et chez les adultes (Lepoutre *et al.*, 2008).

Ce type de vaccin, malgré sa forte immunogénicité, présente néanmoins des inconvénients. Tout d'abord, la protection est restreinte à un certain nombre de sérotypes. De plus, la fabrication de vaccins glycoconjugués est très coûteuse et peu remboursée, rendant le plan de vaccination difficile à mettre en œuvre particulièrement dans les pays en voie de développement. En outre, même si ce vaccin a dramatiquement réduit l'incidence des maladies invasives, l'augmentation des bactériémies et méningites dues à des sérotypes non-inclus dans ce vaccin heptavalent a été reportée (Lepoutre *et al.*, 2008 ; Hicks *et al.*, 2007). À long terme, la distribution de ce vaccin pourrait donc conduire à l'émergence dangereuse des sérotypes invasifs non-inclus dans le vaccin. Par conséquent, une alternative est aujourd'hui à l'étude ; celle d'utiliser des protéines, communes à tous les sérotypes et contribuant à la virulence du pneumocoque, comme vaccin. Ces protéines seraient plus facilement purifiables que les polysaccharides ce qui réduirait le coût d'un éventuel vaccin. Plusieurs protéines candidates ont été identifiées, comme la pneumolysine (la partie non toxique), l'adhésine CbpA, l'autolysine PspA ou encore la neuramidase NanA. Il a été montré que l'immunisation avec ces protéines permettait de protéger contre la pneumonie et la bactériémie (Briles *et al.*, 2003 ; Ogunniyi *et al.*, 2007).

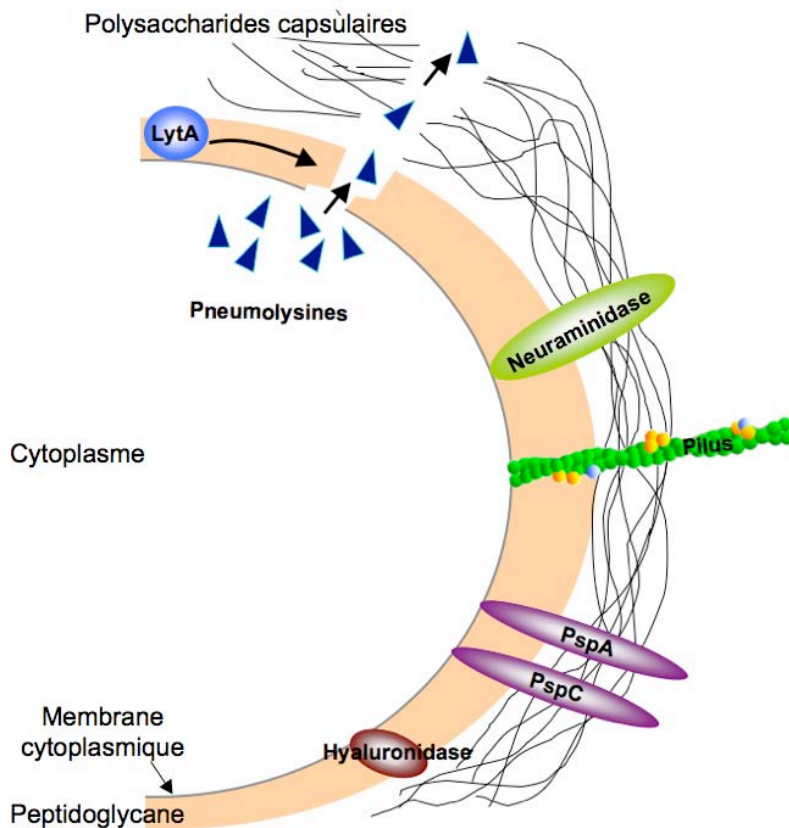
I.3.2. Les antibiotiques

Le 3 septembre 1928, Alexander Flemming découvrit qu'une substance secrétée par le champignon *Penicilium notatum* inhibe la croissance de souche de staphylocoques ; il l'appella aussitôt la pénicilline. Suite à la découverte d'une souche plus productrice (*Penicillium chrysogenum*), la production de la pénicilline a pu être réalisée à échelle industrielle en 1945. La pénicilline fut ainsi très vite mise à profit pour guérir les malades victimes d'infections bactériennes. Les années suivantes ont été considérées comme les années fastes de l'antibiothérapie (1950-1970) avec l'espoir de voir la fin du règne des maladies infectieuses.

Les premiers signes de résistance aux antibiotiques chez le pneumocoque apparurent en 1965 à Boston, puis une souche résistante à la pénicilline fut décrite en 1967 en Australie et enfin une souche multirésistante fut découverte en 1977 en Afrique du Sud (Jacobs *et al.*, 1978). La résistance à la pénicilline se propagea alors rapidement à travers le monde dans les années 1980, particulièrement en Afrique du Sud, en Espagne, en Hongrie et en France, pour laquelle la diminution de la sensibilité à la pénicilline G est apparue en 1979 (Casal, 1982 ; Geslin *et al.*, 1992 ; Marton *et al.*, 1991). En France, le Centre National de Référence des Pneumocoques (CNRP) permet le suivi des sérotypes et de la sensibilité aux antibiotiques du pneumocoque dans le temps. En 2005, le CNRP a reporté que plus de 30 % des souches pneumococciques étaient de sensibilité diminuée à la pénicilline et qu'environ 4 % y étaient résistantes. L'existence de souches ayant une sensibilité moindre à d'autres antibiotiques de la classe des β -lactamines, comme l'amoxicilline et le céfotaxime, a été également reportée. De plus, un taux de résistance aux macrolides de plus de 40 % a été observé (CNRP, rapport d'activité 2006). Le pneumocoque a également acquis des résistances à plusieurs autres familles d'antibiotiques comme les sulfamides, la tétracycline, le chloramphénicol.

L'émergence et la diffusion de la résistance aux antibiotiques chez *Streptococcus pneumoniae* constituent un problème majeur de santé publique. De plus, la couverture et l'efficacité restreintes des vaccins actuellement disponibles incite à rechercher de nouvelles cibles thérapeutiques. Ainsi, l'identification de nouveaux déterminants de la virulence peut faciliter le développement de nouveaux traitements, qu'ils soient au niveau de l'antibiothérapie ou de la vaccination.

II. Les facteurs de virulence du pneumocoque



La virulence est liée aux caractéristiques inhérentes de la bactérie et réside dans la capacité d'une souche à échapper aux systèmes de défense de l'hôte et à se multiplier chez celui-ci. En revanche, le pouvoir pathogène d'une souche bactérienne s'exprime par la création de lésions tissulaires caractéristiques suite à la réaction inflammatoire engendrée par la libération et à l'activation de différents composants bactériens. (AlonsoDeVelasco *et al.*, 1995)

Figure 3 : Représentation schématique des principaux facteurs de virulence du pneumocoque.

La connaissance des facteurs de virulence de *S. pneumoniae* et leur action représente une voie prometteuse pour le développement de nouveaux traitements. Par conséquent, un effort important a été réalisé ces dernières années pour les identifier et les caractériser. 387 gènes pneumococciques impliqués dans la virulence ont ainsi été récemment mis en avant, sans pour autant que toutes leurs fonctions soient déterminées (Hava and Camilli, 2002). La figure 3 représente les facteurs de virulence du pneumocoque dont nous allons discuter le rôle ci-après.

II.1. Les polysaccharides capsulaires

Outre son caractère immunogène qui est à l'origine de la production d'anticorps protecteurs spécifiques chez l'hôte, la capsule polysaccharidique est un déterminant essentiel à la virulence du pneumocoque ; sans capsule, la dose létale augmente d'un facteur supérieur à 10^6 lors d'infections expérimentales (Watson and Musher, 1990). La capsule protège la bactérie de la phagocytose des macrophages et des neutrophiles, permettant ainsi la colonisation *in vivo*. De plus, elle protège certaines protéines de surfaces (telles PspA ou CbpA) des anticorps circulants de l'hôte, leur évitant ainsi une inhibition fonctionnelle (Watson and Musher, 1990). L'épaisseur de la capsule ainsi que sa composition déterminent, à divers degrés, l'habileté d'un sérotype à survivre dans le sang et à causer une infection invasive (Kelly *et al.*, 1994).

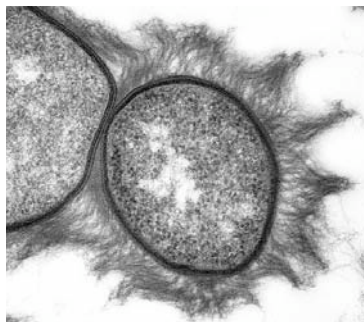


Figure 4 : Vue de la capsule polysaccharidique de la souche TIGR4 de *S. pneumoniae*. (<http://www.rockefeller.edu>).

II.2. La paroi bactérienne

La paroi est composée du peptidoglycane, d'acides teichoïques (polymères contenant de la phosphorylcholine et liés au peptidoglycane) et d'acides lipoteichoïques (polymères contenant de la phosphorylcholine mais ancrés à un glycolipide de la membrane cytoplasmique). Ces deux derniers sont de puissants inducteurs de la réaction inflammatoire capable de provoquer des dommages tissulaires irréversibles (Varon, 2001). La paroi est également impliquée dans l'attachement de pneumocoques non encapsulés sur les cellules endothéliales humaines (Geelen *et al.*, 1993).

II.3. Les protéines pneumococciques

Le pneumocoque possède de nombreuses protéines jouant un rôle dans la pathogénie, que ce soit en permettant l'adhésion, en agissant directement sur les tissus de l'hôte ou encore en médiant l'inflammation. Ces principales protéines sont la protéine A de surface du pneumocoque (PspA), l'adhésine CbpA, la pneumolysine, la neuraminidase, la hyaluronidase et l'autolysine.

II.3.1. La protéine de surface PspA

PspA est une protéine produite par toutes les souches pneumococciques qui inhibe l'activité du complément de l'hôte. Le pneumocoque évite ainsi les mécanismes de défense humains via une diminution du mécanisme d'opsonophagocytose (phagocytose par les macrophages ou neutrophiles facilitée par les anticorps qui recouvrent la cible) (Briles *et al.*, 1988). De plus, PspA joue un rôle dans l'apport en fer pour la croissance *in vivo* du pneumocoque. En effet, elle se lie et inhibe la lactoferrine humaine qui est une protéine ayant un rôle dans la séquestration du fer (Hammerschmidt *et al.*, 1999). Des souches pneumococciques, dans laquelle le gène codant pour PspA a été délété, sont significativement moins virulentes comparées aux souches sauvages (Berry and Paton, 2000).

II.3.2. L'adhésine CbpA

Des souches mutées de *S. pneumoniae* ne produisant plus la protéine CbpA présentent une diminution de leur capacité à coloniser le nasopharynx, à provoquer une pneumonie et une bactériémie dans le modèle murin (Balachandran *et al.*, 2002). Ces observations indiquent que CbpA est un important facteur de virulence, même si le mécanisme d'action de cette protéine est encore pauvrement compris. CbpA jouerait le rôle d'une adhésine de surface et serait impliquée dans l'adhésion aux cellules de la muqueuse nasopharyngée et aux cellules endothéliales pulmonaires. C'est en interagissant avec le récepteur aux immunoglobulines IgR, présent à la surface cellulaire de l'hôte, que CbpA interviendrait dans l'adhésion du pathogène (Zhang *et al.*, 2000).

II.3.3. La pneumolysine

La pneumolysine est une cytotoxine initialement présente dans le cytoplasme de toutes les souches pneumococciques et excrétée lors de leur lyse. Des monomères de pneumolysine se fixent sur la membrane de la cellule cible par l'intermédiaire du cholestérol, s'insèrent puis forment des oligomères intra-membranaires qui s'organisent pour former des pores transmembranaires. Ceci autorise alors l'entrée d'eau dans la cellule qui conduit, le plus souvent, à la lyse cellulaire (Andrew *et al.*, 1997). La dissémination du pneumocoque se retrouve ainsi facilitée (Rubins *et al.*, 1992 ; Rayner *et al.*, 1995). En dehors de son activité toxique, la pneumolysine est également capable d'inhiber la clairance bactérienne des poumons en supprimant le mouvement ciliaire, augmentant, de cette façon, l'adhérence du pneumocoque à l'épithélium bronchial (Steinfort *et al.*, 1989). De plus, la pneumolysine inhibe également les défenses immunes de l'hôte et facilite ainsi la multiplication du pneumocoque dans les poumons et son invasion dans le sang (Paton and Ferrante, 1983).

II.3.4. La neuraminidase

Cette enzyme, présente chez tous les isolats cliniques du pneumocoque, clive l'acide sialique contenu dans les glycolipides, les glycoprotéines et les oligosaccharides présents à la surface des cellules hôtes. Il en résulte alors des dommages cellulaires mais également une exposition de récepteurs utilisés par le pneumocoque pour son adhésion (Paton and Ferrante, 1983). Une diminution de la colonisation du nasopharynx est observée avec des souches de pneumocoque déficientes en neuraminidase (Winter *et al.*, 1997).

II.3.5. La hyaluronidase

Cette enzyme est retrouvée dans la majorité des isolats de *S.pneumoniae* et contribue à la virulence de celui-ci en dégradant l'acide hyaluronique, composant important de la matrice extracellulaire. La migration de la bactérie du site de colonisation vers le système sanguin est, de ce fait, facilitée (Paton and Ferrante, 1983).

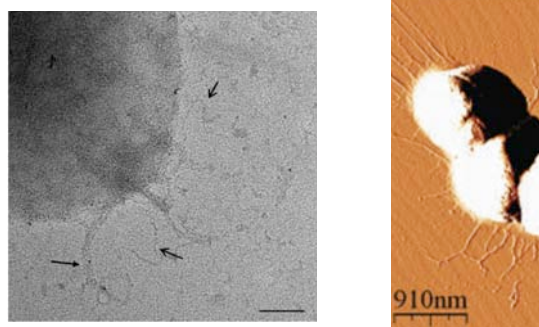
II.3.6. L'autolysine *LytA*

LytA est une enzyme intervenant dans la dégénérescence de la paroi bactérienne en hydrolysant spécifiquement le peptidoglycane (Mosser and Tomasz, 1970). Elle est ainsi impliquée dans l'autolyse cellulaire, à travers laquelle des substances toxiques sont relarguées, comme la pneumolysine, la neuramidase ou encore des produits de dégradation du peptidoglycane qui sont de puissants stimulants de la réaction inflammatoire (Berry and Paton, 2000). *LytA* joue un rôle important dans la virulence du pneumocoque. En effet, une souche déficiente en cette protéine possède un sévère défaut de virulence dans le modèle murin (Berry *et al.*, 1989).

II.4. Les pili

Ces fibres protéiques, qui seront décrites plus en détail dans le paragraphe suivant, sont présentes dans 30% des souches pneumococciques et constituent un facteur de virulence qui n'a été découvert que récemment (Basset *et al.*, 2007 ; Barocchi *et al.*, 2006). La figure 5 met en avant la présence de ces structures filamenteuses à la surface du pneumocoque.

Figure 5 : Microscopie électronique de pneumocoques piliés (Hilleringmann *et al.*, 2008 ; Falker *et al.*, 2008).



Du fait de leur récente découverte, les études concernant leur rôle dans la pathogénie sont encore restreintes aujourd'hui. Néanmoins, *in vitro*, il a été montré que les pili participent au processus d'adhésion du pneumocoque sur les cellules épithéliales pulmonaires (Barocchi *et al.*, 2006). Il a été également reporté que les pili jouent un rôle dans la colonisation et l'invasion des poumons dans le modèle murin (Barocchi *et al.*, 2006 ; Rosch *et al.*, 2008).

D'autre part, il a été montré que les pili augmentent la réponse inflammatoire de l'hôte. En effet, les souches piliées élicitent une réponse immunitaire élevée à cytokines (protéine

impliquée dans la communication des cellules immunitaires) comparée aux souches non piliées (Barocchi *et al.*, 2006). L'utilisation potentielle des pili pneumococciques comme vaccin candidat a été étudiée par Gianfaldoni et ses collègues qui ont montré que les sous-unités protéiques du pilus constituent des antigènes protecteurs (Gianfaldoni *et al.*, 2007). Du fait que les pili ne se retrouvent pas dans l'ensemble des souches pneumococciques, un vaccin combinant les protéines du pilus avec d'autres antigènes serait souhaitable.

III. Les pili, organelles de surface bactériens

L'adhésion des bactéries sur les cellules et tissus hôtes, ainsi que sur les composants de la matrice extracellulaire, constitue une des étapes cruciales dans le processus infectieux. Les adhésines, protéines localisées à la surface bactérienne, sont responsables de cet attachement. Du fait de leurs charges nettes négatives, la bactérie et la cellule hôte se répulsent naturellement rendant difficile l'adhésion du pathogène. Afin de contourner ce problème, ce dernier a mis en place un système permettant à l'adhésine d'être située à distance de la surface bactérienne. Il s'agit d'une structure protéique appelée pili ou fimbriae, pérित्रiche, non flagellaire et filamenteuse, qui s'étend hors de la paroi et sur laquelle l'adhésine est localisée.

Les pili sont composés d'unités protéiques nommées sous-unités pilines. Bien qu'ils aient été décrits comme des organelles de surface jouant un rôle dans l'adhésion, ils sont également impliqués dans d'autres fonctions comme le transfert d'ADN, la formation de biofilm, la motilité de type « twitching » ou encore l'agrégation cellulaire (Koebnik, 2001 ; Manetti *et al.*, 2007 ; Edwards *et al.*, 2008 ; Mattick, 2002).

Les pili sont présents chez les bactéries à Gram-négatif et à Gram-positif. Cependant, leurs études chez l'un et l'autre de ces microorganismes ne sont pas équivalentes. En effet, les pili des bactéries à Gram-négatif ont été découverts les premiers et restent encore aujourd'hui les mieux décrits et compris. En revanche, peu de choses sont connues sur les pili des bactéries à Gram-positif du fait de leur récente découverte, bien que leur intérêt suscite, de nos jours, de nombreuses recherches.

III.1. Les pili des bactéries à Gram-négatif

Les pili des bactéries à Gram-négatif ont été découverts à la fin des années 1940. Depuis, ces structures ont fait l'objet d'intenses recherches et leurs structures, leurs assemblages, leurs régulations et leurs rôles dans la pathogénie sont bien compris.

Classés suivant leur voie d'assemblage, 4 groupes distincts de pili existent : les pili assemblés par la voie « chaperonne/usher », les pili de type IV, les pili curli et les pilis assemblés par la voie chaperonne/usher » alternative (Soto and Hultgren, 1999). Les pili les plus étudiés et les mieux compris sont ceux qui sont assemblés par la voie

« chaperonne/usher ». Les gènes requis pour l'assemblage des pili par cette voie sont généralement organisés en opéron. Durant la formation du pilus, les sous-unités pilines sont tout d'abord secrétées dans le périplasma par la voie générale de sécrétion (système Sec). Elles sont ensuite prises en charge par une chaperonne spécifique, évitant ainsi l'assemblage prématuré du pilus dans le périplasma et l'agrégation des sous-unités pilines. Une protéine, localisée dans la membrane externe bactérienne et appelée « usher », sert ensuite de transporteur pour mener la sous-unité piline à l'extérieur de la bactérie et constitue ainsi une plateforme pour l'assemblage du pilus (Sauer *et al.*, 2004). L'assemblage des pili de ce type est schématisé figure 6.

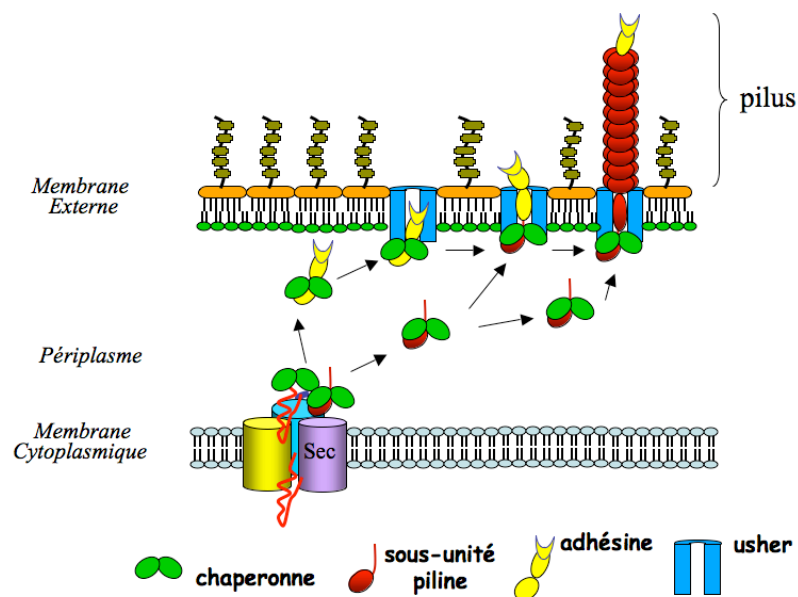


Figure 6 : Assemblage du pilus par la voie « chaperonne/usher » chez les bactéries à Gram-négatif, d'après (Sauer *et al.*, 2004). Sec indique le système de sécrétion.

Une des caractéristiques principales de ces pili est qu'ils sont composés de protéines qui sont liées non-covalamment entre elles. De plus, l'adhésine est retrouvée au sommet du pilus. L'assemblage du pilus chez les bactéries à Gram-négatif ne nécessite pas d'activité enzymatique. En effet, la cohésion des sous-unités pilines entre elles se fait par un mécanisme d'échange de brin donneur. Les sous-unités pilines possèdent une structure apparentée à celle des immunoglobulines (Ig) mais incomplète car un des brins β est manquant. La chaperonne, par le biais d'un de ses brins β , vient compléter le brin manquant de la sous-unité piline. Ensuite, une autre sous-unité piline vient déplacer le brin donneur de la chaperonne et compléter le repliement de type Ig en apportant son extension N-terminale (Scott and Zahner, 2006).

III.2. Les pili des bactéries à Gram-positif

Les premiers pili des bactéries à Gram-positif, détectés par microscopie électronique, ont été décrits en 1968 à la surface du pathogène *Corynebacterium renale* (Yanagawa *et al.*, 1968). Depuis, la présence de pili a été découverte dans d'autres microorganismes : *Streptococcus sanguis* (Fives-Taylor and Thompson, 1985), *Actinomyces naehundii* (Cisar *et al.*, 1988), *Streptococcus salivarius* (Levesque *et al.*, 2001) ou encore *Corynebacterium diptheriae* (Ton-That and Schneewind, 2003). Récemment, des pili ont été également caractérisés chez les 3 principaux streptocoques pathogènes pour l'homme : *Streptococcus pyogenes* (Mora *et al.*, 2005), *Streptococcus agalactiae* (Lauer *et al.*, 2005) et *Streptococcus pneumoniae* (Barocchi *et al.*, 2006). Les pili des bactéries à Gram-positif jouent un rôle majeur dans l'interaction hôte/pathogène mais également dans la colonisation des tissus.

III.2.1. Structure générale

Deux types de pili ont été identifiés par microscopie électronique chez les bactéries à Gram-positif. Dans un premier cas, de courtes et fines « tiges », de longueur comprise entre 70 et 500 nm et de diamètre de 1 à 2 nm ont été localisées à la surface de *S. gordonii*, *S. oralis* et *S. sanguis* (McNab *et al.*, 1999 ; Willcox and Drucker, 1989 ; Willcox *et al.*, 1989). En revanche, des pili plus longs, de 0,3 à 3 µm, flexible, de diamètre compris entre 3 et 10 nm, ont été décrits chez les corynebactéries et chez les 3 streptocoques *S. pyogenes*, *S. agalactiae* et *S. pneumoniae*. La figure 7 illustre quelques différences retrouvées entre les pili des bactéries à Gram-négatif et ceux des bactéries à Gram-positif.

Contrairement aux pili des bactéries à Gram-négatif, ceux des bactéries à Gram-positif sont composés de sous-unités pilines liées covalamment entre elles (Ton-That and Schneewind, 2003, 2004). Ces sous-unités ne peuvent donc pas être dissociées les unes des autres par chauffage ou par l'agent SDS. Cette propriété est notamment utilisée pour déterminer si une protéine fait partie du pilus. Une simple analyse en gel SDS permet alors d'étudier la présence de pilus, caractérisée par des bandes protéiques de haut poids moléculaire.

Le corps du pilus est composé de la polymérisation d'une même sous-unité, appelée

piline majeure. Des **pilines mineures**, ou **auxiliaires**, sont également présentes sur le corps du pilus (au sommet, à la base ou de part et d'autre de la structure du pilus) mais ne sont pas requises pour l'intégrité du pilus (Ton-That and Schneewind, 2003, 2004). Les pili sont ancrés covalamment au peptidoglycane de la bactérie.


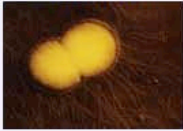
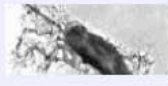


	Bactéries à Gram-négatif			Bactéries à Gram-positif	
	Pili de type I	Pili de type IV	Pili curli	Fibrille	Pili
Exemples d'espèces	<i>E.coli</i> entéropathogène	<i>E.coli</i> , <i>Neisseria</i> , <i>Pseudomonas</i>	<i>E.coli</i> (quelques souches)	<i>S.salivarius</i>	<i>S. pneumoniae</i> , <i>Corynebacterium</i>
Microscopie Électronique					
Morphologie	Tige rigide avec une adhésine flexible au sommet	Tige flexible	Filaments en coiled-coil	Courte, tige fine	Tige flexible
Longueur	1-2 µm	1-2 µm	1-2 µm	0,07-0,5 µm	0,3-03 µm
Diamètre	7 nm (Sommet : 2-3 nm)	5-6 nm	1-2 nm	1-2 nm	3-10 nm
Biogenèse du pilus	Voie Chaperonne/usher	Système de sécrétion du type II	Extracellulaire	ND	Activité de sortases
Récepteurs cellulaires des adhésines	Glycoprotéines, Glycolipides	Glycolipides, CD46, C4BP	Fibronectine, laminine, plasminogène	Fibronectine	Collagène Fibronectine
Organisation génomique	Opéron	Ilôt de pathogénicité	Opéron	ND	Ilôt de pathogénicité
Immunogénicité	Protection engendrée par les protéines du pilus purifiées additionnées d'un adjuvant	Protection engendrée par les protéines du pilus purifiées	ND	ND	Protection engendrée par les protéines du pilus purifiées
Fonctions	Adhésion sur les tissus hôtes, co-aggrégation, immunomodulation, formation de biofilm	Adhésion sur les tissus hôtes, co-aggrégation, immunomodulation, motilité, transfert d'ADN	Adhésion sur les tissus hôtes, co-aggrégation, immunomodulation, formation de biofilm	Adhésion sur les tissus hôte, co-aggrégation,	Adhésion sur les tissus hôte, co-aggrégation, immunomodulation,

Figure 7 : Types de pili bactériens d'après (Telford *et al.*, 2006). ND signifie non déterminé.

L'immunomodulation réfère à la capacité de la bactérie à modifier l'intensité des réactions immunitaires de l'hôte.

III.2.2. Les sortases, enzymes catalysant la formation du pilus

Chez les bactéries à Gram-positif, la biogenèse du pilus nécessite une activité enzymatique qui est réalisée par les sortases et qui va permettre de lier de manière covalente les sous-unités pilines entre elles. Une ou plusieurs sortases peuvent être impliquées dans la formation du pilus. Nous discuterons en détail de ces sortases, ainsi que du mécanisme par lequel elles assemblent le pilus, dans la partie IV de ce manuscrit.

III.2.3. Les pilines majeures et auxiliaires

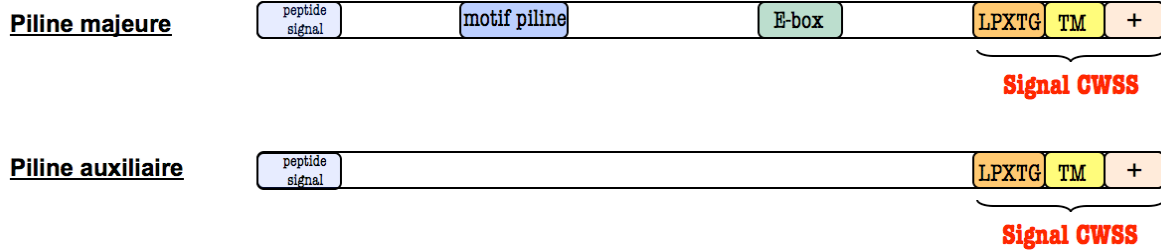
La polymérisation de la piline majeure en pilus est nécessaire pour que s'insèrent dessus les protéines auxiliaires ; sans polymérisation, ces dernières restent ancrés sur le peptidoglycane. Toutes ces protéines possèdent un motif bien particulier, appelé **motif CWSS** (Cell Wall Sorting Signal) (figure 8A). Ce signal, localisé en C-terminal, comprend :

- une séquence peptidique appelé **motif LPXTG**
- un domaine hydrophobe, de longueur variable (15 à 30 résidus)
- une queue chargée positivement, de 4 à 9 résidus dont au moins une arginine.

Il a été montré que le motif CWSS d'une protéine est indispensable pour son ancrage sur le peptidoglycane bactérien (Schneewind *et al.*, 1992). Ainsi, la queue chargée positivement évite l'excrétion totale de la protéine et le domaine hydrophobe permet l'ancrage de la protéine à la membrane cytoplasmique (figure 8B). Le motif LPXTG, quant à lui, permet la reconnaissance de la protéine par les sortases provoquant son attachement sur les autres protéines du pilus et/ou sur le peptidoglycane. En absence de motif LPXTG, la protéine reste ancrée sur la membrane (Schneewind *et al.*, 1993).

Les sous-unités pilines possèdent également un peptide signal situé du côté N-terminal leur permettant d'être exportée via le système Sec. Ce peptide signal est composé de 1 à 3 résidus chargé positivement (domaine N), suivis d'une région hydrophobique de 10 à 15 résidus (domaine H) puis d'une région polaire qui constitue le site de clivage (domaine C). La sous-unité piline est synthétisée dans le cytoplasme bactérien et va être transloquée à travers la membrane cytoplasmique grâce à la machinerie cellulaire Sec (figure 8B). Cette dernière est composée d'un complexe protéique membranaire (SecYEG) et d'une protéine périphérique (SecA) fournissant l'énergie nécessaire au transport. Une signal-peptidase, liée à la membrane cytoplasmique, intervient pour cliver le peptide signal de la protéine précurseur dans le périplasme (Driessen and Nouwenn, 2008).

A



B

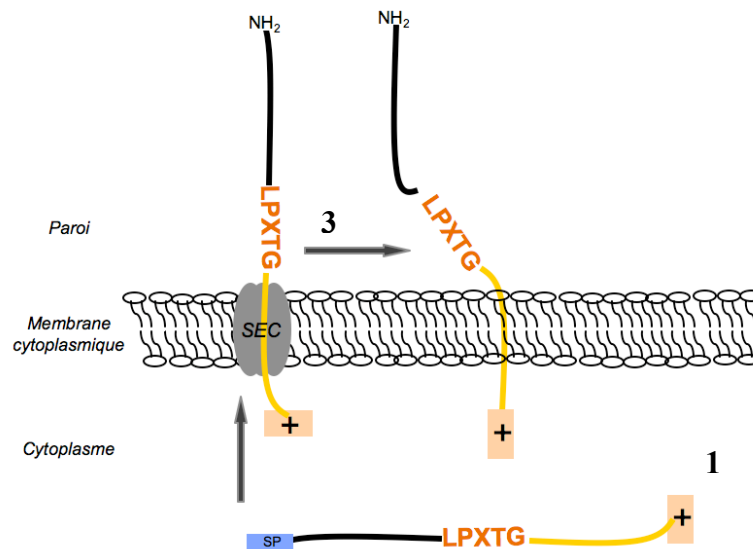


Figure 8 : Les pilines majeures et auxiliaires des pili des bactéries à Gram-positif. A/ Topologie des protéines majeures et auxiliaires. Le motif CWSS est composé du motif LPXTG, d'un domaine transmembranaire (TM) et d'une queue chargée positivement (+). B/ Ancrage des protéines portant un signal CWSS à la membrane bactérienne. La protéine produite dans le cytoplasme (1) va être transloquée via le système Sec (2) et être ancrée sur la membrane à travers son domaine transmembranaire (3).

Toutefois, deux autres motifs sont également distinguables et importants pour l'attachement des protéines du pilus entre elles. Il s'agit du **motif piline** et du **motif E-Box** (Ton-That and Schneewind, 2003 ; Ton-That *et al.*, 2004) (figure 8A). Le motif piline, composé de la séquence peptidique WXXXVXVYPKN, n'est présent et conservé que dans les protéines majeures. La lysine de ce motif joue un rôle crucial : le remplacement de cette lysine dans la protéine majeure de *C. diptheriae* abolit la polymérisation du corps du pilus (Ton-That *et al.*, 2004). Cette lysine est pontée à la thréonine du motif LPXTG d'une sous-unité suivante, via l'activité sortasique, et permet ainsi à deux pilines majeures d'être liée covalamment entre elles et de constituer ainsi, au fur et à mesure, le corps du pilus. Le motif E-Box, dont la séquence peptidique YXLXETXAPXGY comporte un glutamate hautement

conservé, n'est également présent que dans les pilines majeures. Il a été suggéré que ce motif jouait un rôle dans l'attachement des pilines auxiliaires sur le pilus, bien que le mécanisme exact reste encore inconnu. La mutation du glutamate du motif E-Box dans la piline majeure de *C. diptheriae* abolit effectivement l'incorporation de la piline mineure sur la structure du pilus (Ton-That *et al.*, 2004). Il est à noter que les sous-unités auxiliaires peuvent éventuellement se retrouver localisées sur le peptidoglycane et non sur le pilus, comme cela est le cas chez *S. agalactiae* (Krishnan *et al.*, 2007). Bien que le rôle des pilines mineures n'ait pas été entièrement analysé, il semblerait qu'elles puissent jouer le rôle d'adhésine (Nelson *et al.*, 2007 ; Mandlik *et al.*, 2007 ; Krishnan *et al.*, 2007).

III.2.4. Quelques exemples de pili de bactéries à Gram-positif

III.2.4.1. *Corynebacterium diptheriae*

C. diptheriae, agent causal de la diphtérie, est un des premiers organismes où la présence de pili a été décrite (Yanagawa and Honda, 1976). Trois types de pili sont présents à sa surface : pili SpaA, pili SpaD et pili SpaH (pour Sortases-mediated pili assembly). Ces pili sont nommés suivant la piline majeure qui les structure. Ainsi, les pili SpaA sont composés de la protéine SpaA qui est polymérisée pour former le corps du pilus et de pilines auxiliaires SpaB, localisée à intervalles réguliers le long du pilus, et SpaC, située au sommet du pilus (Ton-That and Schneewind, 2003). Trois clusters qui incluent les gènes codant pour les sous-unités pilines et les sortases sont présents chez *C. diptheriae* (figure 9).

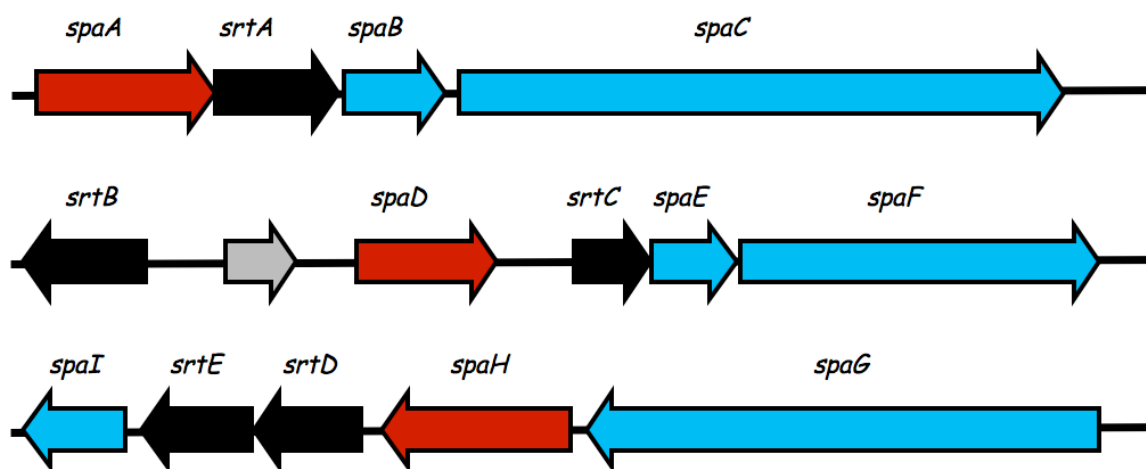


Figure 9 : Opérons groupant les gènes du pilus chez *C. diptheriae*, d'après (Mandlik *et al.*, 2007). Les gènes codant pour les sortases sont représentés en noir, ceux codant pour les pilines majeures sont en rouge et ceux codant pour les pilines mineures sont en bleu. En gris, est représenté un gène non caractérisé.

La présence de multiples sortases pose la question de la spécificité de substrats. Le rôle de certaines sortases de *C. diphtheriae* a été ainsi investi. La sortase A (SrtA) a été montrée comme essentielle à l'assemblage des pili de type SpaA ; SrtA polymérise la piline majeure SpaA en pilus (Ton-That and Schneewind, 2003). Au niveau du second opéron, aussi bien SrtB que SrtC sont capables de polymériser la piline majeure SpaD en pilus. Cependant, seule SrtB est requise pour l'incorporation de la piline auxiliaire SpaE sur le pilus (Gaspar and Ton-That, 2006). Un gène additionnel codant pour la sortase F (ou SrtF) est localisé dans le génome en dehors des 3 opérons mentionnés ci-dessus. SrtF, quant à elle, est requise pour l'ancrage du pilus sur le peptidoglycane, mais n'intervient pas dans son assemblage (Swaminathan *et al.*, 2007).

Les pili de *C. diphtheriae*, et plus précisément les pilines mineures SpaB et SpaC, sont impliqués dans l'adhérence et la colonisation des cellules humaines pharyngeales (Mandlik *et al.*, 2007).

III.2.4.2. *Streptococcus agalactiae*

Ce streptocoque colonise principalement, de façon asymptomatique, les tractus digestif et génital de l'homme. Cependant, il peut être la cause d'infections invasives, comme la septicémie ou la méningite, dans trois populations distinctes : les nouveau-nés, les femmes enceintes et les personnes de plus de 65 ans ou immunosupprimées.

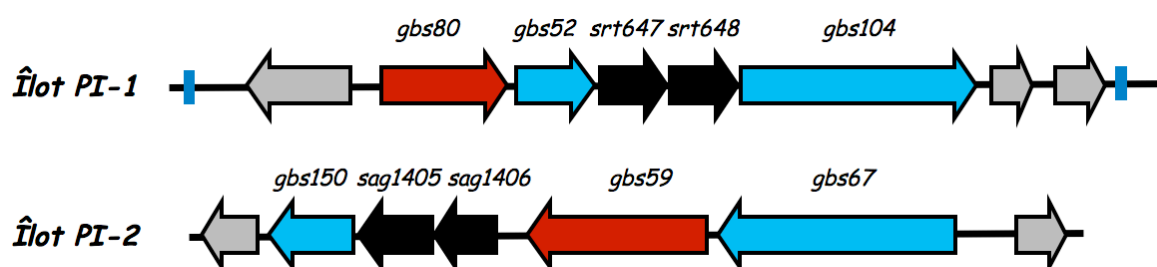


Figure 10 : Îlots de pathogénicité groupant les gènes du pilus chez *S. agalactiae*, d'après (Mandlik *et al.*, 2007). Les gènes codant pour les sortases sont représentés en noir, ceux codant pour les pilines majeures en rouge et ceux codant pour les pilines mineures en bleu. En gris sont représentés les gènes non caractérisés. Les rectangles bleus se réfèrent à des séquences d'insertions.

Deux types de pili, dont les composants sont codés par les îlots PI-1 et PI-2, sont présents à la surface de *S. agalactiae*. Suivant les souches, les deux îlots peuvent être distribués simultanément ou non. L'îlot PI-1, représenté en figure 10, code pour deux sortases et 3 sous-unités pilines. Le corps du pilus est formé par la polymérisation de la piline majeure gbs80 tandis que les protéines gbs52 et gbs104 sont localisées de part et d'autre du pilus (Lauer *et al.*, 2005). En ce qui concerne le second type de pilus, gbs59 est polymérisée pour former le corps du pilus sur lequel est insérée gbs67 et gbs150 (Rosini *et al.*, 2006).

Chez *S. agalactiae* également le rôle des sortases dans la formation du pilus a été étudié. L'une ou l'autre des sortases srt647 et srt648 est requise pour polymériser gbs80. srt647 est, de plus, impliquée dans l'incorporation de gbs52 sur la structure du pilus alors que srt648 est, elle, impliquée dans l'incorporation de gbs104. En ce qui concerne le pilus codé par l'îlot PI-2, les sortases sag1405 et sag1406 sont capables, toutes les deux, de polymériser gbs59 pour former le corps du pilus. Néanmoins, sag1405 est spécifique de l'incorporation de gbs67 sur le pilus et sag1406 spécifique de l'incorporation de gbs150 (Rosini *et al.*, 2006).

Il a été montré que les pili de *S. agalactiae* jouent un rôle dans l'adhérence de la bactérie ainsi que dans son invasion au niveau du cerveau (Maisey *et al.*, 2007). Gbs52 a été la première piline mineure d'un pilus d'une bactérie à Gram-positif dont la structure cristallographique a été résolue (Krishnan *et al.*, 2007). Cette protéine possède deux domaines, N1 et N2, ayant un repliement type immunoglobuline. Chaque domaine comprend 7 brins β antiparallèles qui s'assemblent pour former un tonneau β (figure 11). Le domaine N1 est suffisant pour permettre l'adhésion sur les cellules épithéliales pulmonaires.

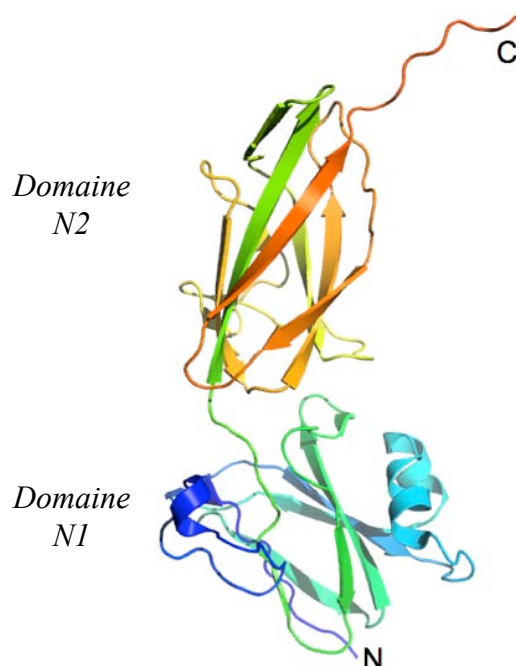


Figure 11 : Structure de la piline mineure gbs52 de *S. agalactiae* (code DPB : 2PZ4). Deux domaines, notés N1 et N2, présentent un repliement similaire à celui des immunoglobulines.

III.2.4.3. *Streptococcus pyogenes*

S. pyogenes est un pathogène strictement humain responsable, dans la majorité des cas, d'infections bénignes et non invasives comme l'angine ou l'impétigo. Néanmoins, cette bactérie peut être responsable d'infections invasives graves telles que les infections cutanées nécrosantes et le syndrome de choc toxique streptococcique.

Le corps du pilus, chez *S. pyogenes*, est formé par la polymérisation d'une même sous-unité dont le gène est hautement variable : *spy0128* (sérotipe M1), Tee6 (sérotipe M6), Orf80 (sérotypes M3, M5, M18,M49) ou Eft LSL.A (sérotipe M12). Une protéine auxiliaire, dont la séquence génomique est également variable, est détectée le long du pilus : *spy0130* (sérotipe M1), *Orf82* (sérotypes M3, M5, M18,M49) ou Orf2 (sérotipe M12).

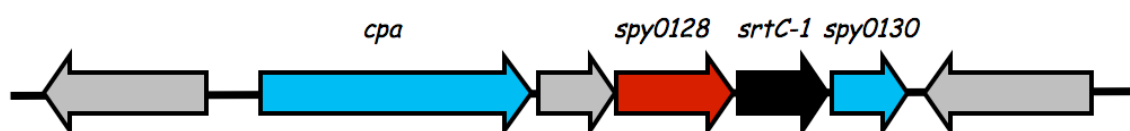


Figure 12 : Opéron regroupant les gènes du pilus chez *S. pyogenes*, d'après (Mandlik *et al.*, 2007). En noir est représenté le gène codant pour la sortase, en rouge celui qui code pour la piline majeure et en bleu les gènes codant pour les pilines mineures. Les gènes non caractérisés sont représentés en gris.

La structure de *spy0128*, représentée en figure 13A, est la seule structure aujourd'hui disponible d'une piline majeure de pilus (Kang *et al.*, 2007). *Spy0128* possède un repliement en deux domaines composés de brins β . Deux ponts intramoléculaires ont été reportés dans cette structure ; ceux ci sont formés entre le groupe amine de la chaîne latérale d'une lysine et le groupement carboxyle de la chaîne latérale d'une asparagine, et ceci dans un processus catalytique impliquant un résidu glutamate situé à proximité (figure 13B). Ces ponts sont supposés jouer un rôle dans la stabilité de la protéine.

Il a été reporté que les pili de *S. pyogenes* contribuent à l'adhérence du pathogène sur les cellules pharyngeales de l'hôte mais également à la formation de biofilm (Manetti *et al.*, 2007).

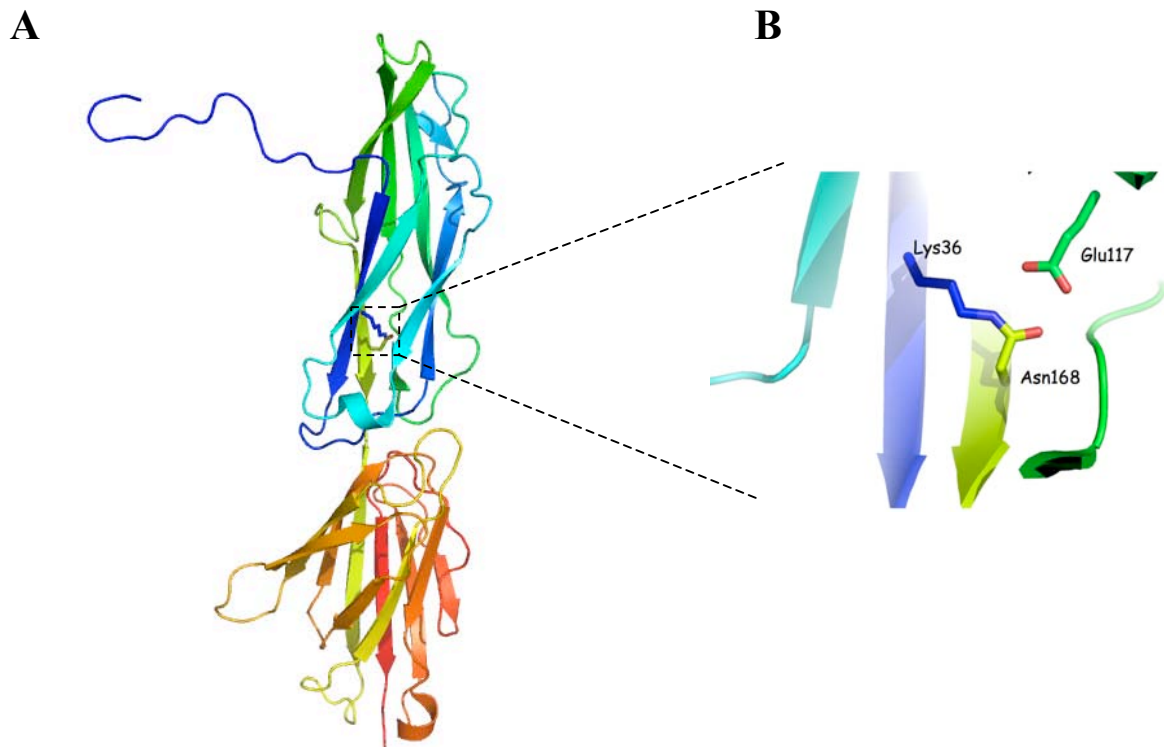


Figure 13 : Structure de la piline majeure spy0128 de *S. pyogenes* (code PDB : 3B2M). A/ spy0128 est composée de deux domaines constitués de brins β . B/ Pont intramoléculaire de spy0128.

III.2.4.4. *Streptococcus pneumoniae*

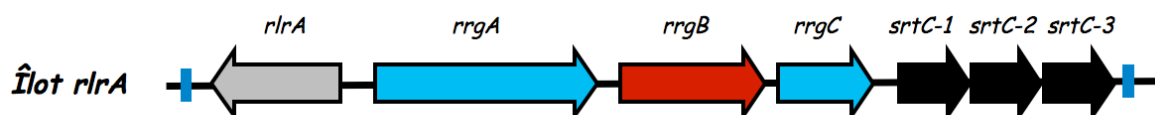
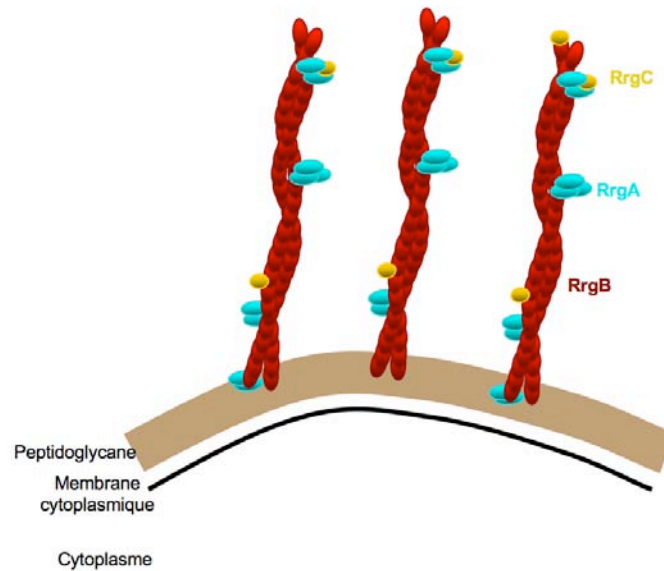


Figure 14 : Îlot de pathogénicité *rlrA* de *S. pneumoniae* d'après (Mandlik *et al.*, 2007). Les gènes codant pour les sortases sont représentés en noir, ceux codant pour les pilines mineures en bleu et celui codant pour la piline majeure en rouge. Le gène *rlrA* code pour un régulateur positif. Les rectangles bleus indiquent la présence d'éléments d'insertion.

Les gènes requis pour la production des pili chez *S. pneumoniae* sont localisés dans l'îlot de pathogénicité *rlrA* (figure 14), qui comprend 6 gènes codant pour 3 sortases (*SrtC-1*, *SrtC-2*, *SrtC-3*) et 3 protéines structurales (*RrgA*, *RrgB*, *RrgC*). Le gène *rlrA* code pour un régulateur positif des 6 gènes voisins. Cet îlot comporte deux séquences d'insertion indiquant qu'il est mobile ; cela est en accord avec le fait que les pili ne sont présents que dans 30 % des

souches de pneumocoques. Le corps du pilus est formé par la polymérisation de la piline majeure RrgB. RrgA et RrgC se localisent ponctuellement le long du pilus (Hilleringmann *et al.*, 2008) (figure 15A). Il a été montré que deux protofilaments s'assemblent en une structure « coiled-coil » pour former le corps du pilus, comme ceci est montré figure 15.

A



B

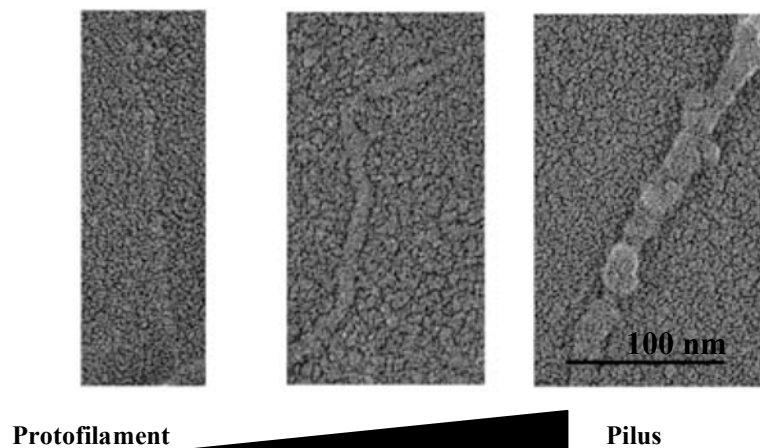


Figure 15 : Composition des pili de *S. pneumoniae*. A/ Schéma des pili pneumococciques se composant des protéines RrgA, RrgB et RrgC. B/ Microscopie électronique de pili du pneumocoque montrant que ceux-ci sont des structures assemblées en « coiled-coil », (Hilleringmann *et al.*, 2008).

RrgA a été décrite comme une adhésine, capable de lier la fibronectine, la laminine et le collagène de type I, qui sont des composants de la matrice extracellulaire (Hilleringmann *et al.*, 2008). De plus, des tests *in vitro* sur des cellules épithéliales pulmonaires montrent que les

souches pneumococciques mutantes, dont le gène codant pour RrgA a été délété, sont moins adhérentes que les souches sauvages (Nelson *et al.*, 2007). Aucun rôle n'a encore été suggéré pour RrgC à ce jour.

Ce n'est que récemment que le rôle des sortases du pneumocoque dans l'assemblage du pilus a été étudié. C'est par des expériences réalisées *in vivo*, en délétant des souches pneumococciques de un ou plusieurs gènes de sortases et en analysant les phénotypes par microscopie électronique et par gel SDS que LeMieux ainsi que Falker et leurs collègues ont décrit le rôle de SrtC-1, SrtC-2 et SrtC-3. L'absence des gènes codant pour ces trois sortases abolit la polymérisation de RrgB, indiquant que une ou plusieurs sortases sont requises pour polymériser RrgB en fibres (LeMieux *et al.*, 2008). En construisant de simples mutants de SrtC-1, SrtC-2 ou SrtC-3, Le Mieux et ses collègues conclurent que la polymérisation de RrgB ne dépend pas que d'une sortase et qu'une certaine redondance d'activité est retrouvée. Cependant, une spécificité concernant l'incorporation des protéines auxiliaires sur le pilus par les différentes sortases a été montrée. Ainsi, SrtC-2 n'est pas requise pour l'incorporation de RrgA ou RrgC sur le pilus alors que SrtC-1 et SrtC-3 sont capables d'incorporer ces deux protéines sur le pilus (LeMieux *et al.*, 2008). Au contraire, Falker et ses collègues ont montré que SrtC-1 et SrtC-2 étaient capables d'incorporer RrgA sur le corps du pilus mais que seule SrtC-1 incorporait RrgC sur le pilus. De plus, ils montrèrent que SrtC-1 étaient la principale enzyme polymérisant RrgB en fibre (SrtC-2 étant capable, mais dans une moindre mesure, de catalyser la formation de la fibre) (Falker *et al.*, 2008). Les divergences retrouvées entre les travaux de LeMieux et Falker évoquent certainement la complexité du processus de formation du pilus chez *S. pneumoniae*.

Toutes les souches de pneumocoque possèdent un gène codant pour la sortase A (SrtA), localisé en dehors de l'îlot *rhrA*. Cette sortase ancre des protéines possédant un motif LPXTG sur le peptidoglycane. De premiers travaux ont montré que SrtA n'intervenait ni dans l'assemblage, ni dans l'accrochage du pilus sur le peptidoglycane (LeMieux *et al.*, 2008).

Récemment, la présence d'un pilus additionnel a été mise en évidence dans 16 % des souches pneumococciques (Bagnoli *et al.*, 2008). L'îlot PI-2, qui réfère à ce pilus, est distribué ou non avec l'îlot *rhrA* et code pour deux protéines structurales (PitA et PitB), une protéine (SipA) ayant une homologie avec les protéines « signal-peptidases » et deux sortases (SrtG-1 et SrtG-3) (figure 16). La polymérisation du pilus requiert la protéine majeure PitB

aussi bien que la sortase SrtG-1 et la protéine SipA (Bagnoli *et al.*, 2008). SipA jouerait le rôle d'une chaperonne prévenant une agrégation prématurée de la sous-unité piline (Zahner and Scott, 2008). D'autre part, il a été montré que le pilus PI-2 est impliqué dans l'adhérence du pneumocoque sur les cellules bronchiales et nasopharyngeales de l'hôte (Bagnoli *et al.*, 2008).

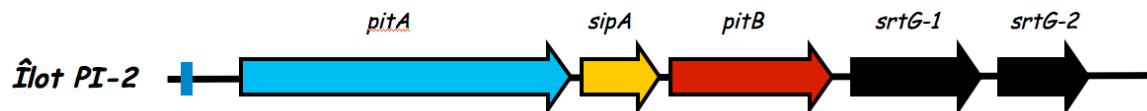


Figure 16 : Îlot PI-2 codant pour un second type de pilus chez *S. pneumoniae*. Les gènes codant pour les sortases sont en noir, le gène codant pour la piline majeure en rouge et celui codant pour la piline mineure en bleu. En orange est indiqué le gène codant pour une protéine ayant une homologie avec les signal-peptidases.

IV. Les sortases et l'art d'accrocher des protéines cibles à la paroi bactérienne

Comme énoncé précédemment, les pili des bactéries à Gram-positif sont assemblés et ancrés sur le peptidoglycane grâce à l'action d'enzymes spécifiques, les sortases. Les sortases ont également la fonction d'ancrer sur la paroi bactérienne des protéines de surface portant un motif LPXTG qui sont des protéines impliquées notamment dans des interactions avec l'hôte et l'environnement.

Les sortases sont des enzymes reconnaissant un motif bien spécifique porté par les protéines cibles et catalysant une réaction de transpeptidation. Les sortases sont des transpeptidases à cystéines. En effet, elles possèdent un mécanisme catalytique impliquant un acide aminé cystéine présent dans leur site catalytique.

La sortase A de *Staphylococcus aureus*, dont la fonction est d'ancrer des protéines cibles sur le peptidoglycane bactérien, a été la première sortase décrite et reste, à ce jour, la sortase la mieux caractérisée biochimiquement et structuralement. Par homologie de séquences, plusieurs homologues ont ensuite été identifiés chez les bactéries à Gram-positif. Une récente étude a permis de classer les sortases en 4 groupes suivant leurs homologies de séquences et leurs fonctions distinctes dans les bactéries à Gram-positif (Dramsı *et al.*, 2005). Plus d'un gène codant pour une sortase par génome est souvent retrouvé (Pallen *et al.*, 2001) (figure 17).

	<div>Sortases de classe A</div> <div>Sortases de classe B</div> <div>Sortases de classe D</div> <div>Sortases de classe C</div>			
<i>Streptococcus pneumoniae</i>	●			● ● ●
<i>Streptococcus pyogenes</i>	●	●		●
<i>Streptococcus agalactiae</i>	●			● ● ● ●
<i>Streptococcus suis</i>	●			● ● ● ●
<i>Streptococcus mutans</i>	●			
<i>Corynebacterium diphtheriae</i>	●			● ● ● ● ●
<i>Corynebacterium jeikeium</i>	●			●
<i>Corynebacterium glutamicum</i>	●			
<i>Enterococcus faecalis</i>	●			●
<i>Enterococcus faecium</i>	●			● ● ● ●
<i>Clostridium perfringens</i>	●	●	●	●
<i>Bacillus cereus</i>	●	●	● ●	●
<i>Bacillus anthracis</i>	●	●	●	
<i>Bacillus subtilis</i>			●	
<i>Lactococcus lactis</i>	●	●		
<i>Listeria monocytogenes</i>	●	●		
<i>Staphylococcus aureus</i>	●	●		
<i>Actinomyces naeslundii</i>	●			● ●
<i>Streptomyces coelicolor</i>			● ●	
<i>Streptomyces avermitilis</i>			● ● ● ●	

Figure 17 : Classes de sortases de diverses bactéries à Gram-positif, d'après (Mandlik *et al.*, 2007) et (Dramsi *et al.*, 2005).

IV.1. Classification des sortases

IV.1.1. Les sortases de classe A (SrtA)

Depuis ces dernières années, plusieurs études ont vu le jour concernant le rôle des sortases de classe A chez différents organismes à Gram-positif, comme *L. monocytogenes* (Bierne *et al.*, 2002), *S. pyogenes* (Barnett and Scott, 2002), *S. pneumoniae* (Kharat and Tomasz, 2003), *S. suis* (Osaki *et al.*, 2002), *S. gordonii* (Bolken *et al.*, 2001), *S. mutans* (Cossart and Jonquieres, 2000) et *S. agalactiae* (Lalioui *et al.*, 2005). Chaque bactérie ne possède qu'un unique gène codant pour une SrtA dans son génome. Toutes les sortases de classe A possèdent, en N-terminal, un peptide signal hydrophobe qui permet leur sécrétion à travers la membrane cytoplasmique et leur ancrage à la membrane. De plus, en C-terminal, se trouve la séquence consensus TLXTC contenant la cystéine catalytique, dans laquelle X est souvent une valine, une isoleucine ou une thréonine.

Les sortases de classe A ancrent sur le peptidoglycane des protéines cibles portant la séquence peptidique LPXTG (Bierne *et al.*, 2002, 2004 ; Mazmanian *et al.*, 2002 ; Osaki *et al.*, 2002). Comme le montre la figure 18, SrtA reconnaît le motif LPXTG de la protéine à ancrer et clive, via sa cystéine nucléophile, la liaison peptidique entre la thréonine et la glycine. Il en résulte alors un intermédiaire acyl-enzyme dans lequel la thréonine du motif LPXTG se retrouve liée à la cystéine de la sortase. L'attaque nucléophile d'un groupement amine de la chaîne peptidique du lipide II permet ensuite de libérer l'enzyme. Le lipide II est le précurseur de la synthèse du peptidoglycane constitué d'un disaccharide (N-acétyl-glucosamine et acide N- acétyl-muramique) sur lequel est fixé une chaîne peptidique et sur laquelle peut également se brancher une chaîne peptidique additionnelle. La protéine cible se retrouve alors branchée sur le lipide II. Ce lipide II va ensuite être utilisé par les Penicillin Binding Proteins pour élaborer le peptidoglycane. De ce fait, la protéine cible se retrouve ancrée sur le peptidoglycane (Schneewind *et al.*, 1995 ; Ton-That *et al.*, 1997 ; Navarre and Schneewind, 1999). Les sortases de classe A sont également appelées sortases « housekeeping ».

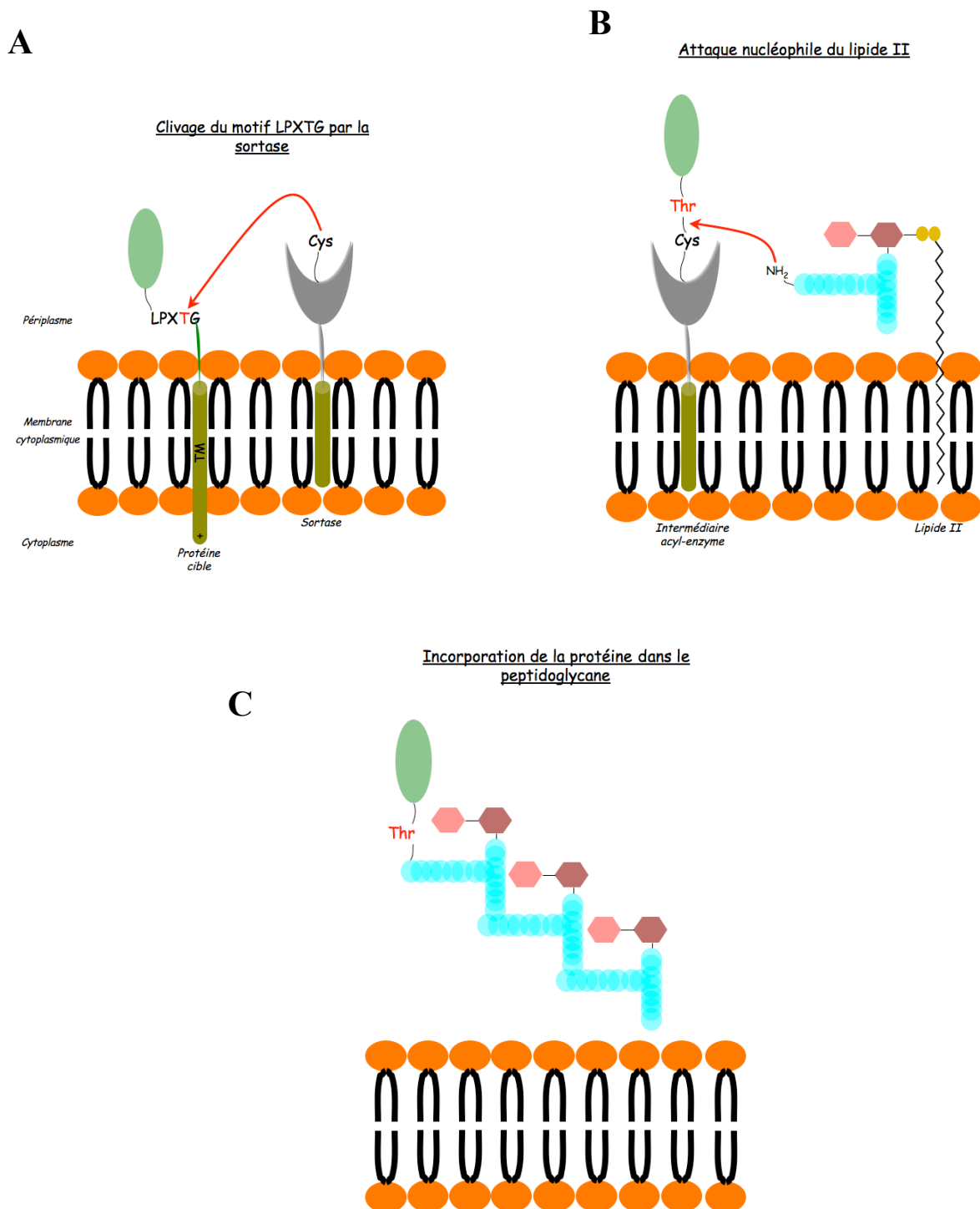


Figure 18 : Mécanisme d'ancrage de protéines possédant un motif LPXTG sur le peptidoglycane bactérien.
 A/ Le motif LPXTG de la protéine ancrée à la membrane cytoplasmique est reconnu puis clivé par la sortase de classe A au niveau de la thréonine. B/ L'intermédiaire acyl-enzyme est ensuite résolu par l'attaque nucléophile du groupement amine de la chaîne polypeptidique du lipide II. Ceci permet de régénérer l'enzyme et de lier covalamment la protéine sur le lipide II. C/ Le lipide II est utilisé dans la synthèse du peptidoglycane par les Penicillin Binding Protein permettant ainsi à la protéine d'être localisée sur le peptidoglycane.

Les sortases de classe A sont responsables de l'ancrage de protéines impliquées dans l'adhésion sur des tissus spécifiques, dans l'invasion des cellules hôtes ou encore dans la résistance à la phagocytose. Une inactivation de SrtA affecte la capacité du pathogène à établir une infection, comme c'est le cas chez *S. aureus* (Mazmanian *et al.*, 2002) ou *L. monocytogenes* (Bierne *et al.*, 2002). Les sortases de classes A ont donc été décrites comme des cibles attractives pour le développement de nouveaux traitements contre les infections bactériennes.

IV.1.2. Les sortases de classe B (SrtB)

Cette classe de sortases est celle qui contient le moins de membres. Elles ont été notamment étudiées chez *S. aureus* (Mazmanian *et al.*, 2002), *B. anthracis* (Maresso *et al.*, 2006) et *L. monocytogenes* (Bierne *et al.*, 2004). Toutes les sortases de classes B possèdent un peptide signal en N-terminal, 3 domaines (B1, B2 et B3) non présents chez les autres types de sortases, et le motif TLXTC en C-terminal dans lequel X est généralement une sérine. Contrairement aux sortases de classe A, les gènes codant pour SrtB et ses potentiels substrats sont localisés dans un même locus.

Les sortases de classe B ancrent sur le peptidoglycane des protéines cibles portant non pas la séquence LPXTG, comme c'est le cas pour les sortases de classe A, mais des protéines possédant le motif NPQTN (Mazmanian *et al.*, 2002). Après reconnaissance du motif NPQTN de la protéine cible, SrtB clive la liaison peptidique entre la thréonine et l'asparagine. La protéine cible se retrouve ensuite accrochée au peptidoglycane via une liaison amide entre la thréonine du motif NPQTN et la chaîne peptidique du lipide II. Le mécanisme par lequel l'ancrage à la paroi bactérienne se fait est le même que celui des sortases de classe A (figure 18).

Les sortases de classe B sont impliquées dans l'ancrage au peptidoglycane de protéines participant à l'acquisition de fer. Le gène codant pour SrtB se localise dans un opéron contenant des protéines ayant un rôle dans l'obtention de ce métal. De plus, les protéines cibles de SrtB possèdent plusieurs domaine répétitifs, appelés NEAR (Near Transporter), supposés jouer un rôle dans la capture du fer (Andrade *et al.*, 2002). A ce jour, le rôle de SrtB a été démontré chez *S. aureus*, *L. monocytogenes* et *B. anthracis* (Mazmanian

et al., 2003 ; Bierne *et al.*, 2004 ; Maresso and Schneewind, 2006 ; Maresso *et al.*, 2006). Des souches, dans lesquelles le gène codant pour SrtB a été délété, possèdent, notamment, un défaut dans leur virulence (Mazmanian *et al.*, 2002 ; Jonsson *et al.*, 2003 ; Weiss *et al.*, 2004).

IV.1.3. Les sortases de classe C (SrtC)

C'est le groupe contenant le plus de membres. Les gènes codant pour les sortases de classe C sont souvent présents en plusieurs copies dans le génome. Ces sortases sont les seules qui possèdent, en C-terminal, un domaine hydrophobe servant d'ancrage transmembranaire et une proline conservée après la séquence consensus TLXTC. Le peptide signal est toujours présent en N-terminal (Dramsi *et al.*, 2005).

Des premiers travaux réalisés sur *A. naeslundii* et *C. diptheriae* ont montré que ce type de sortases serait impliqué dans la formation de pilus (Yeung and Ragsdale, 1997 ; Yeung *et al.*, 1998 ; Ton-That and Schneewind, 2003). La présence de pili découverts à la surface de nombreuses bactéries à Gram-positif possédant des sortases de classe C, ainsi que l'augmentation des études concernant le rôle de ces sortases, ont confirmé l'implication des sortases de classe C dans la biogenèse des pili.

Lors du processus de biogenèse du pilus, trois grandes étapes sont à distinguer : la polymérisation de la piline majeure pour former le corps du pilus, l'ancrage du pilus sur le peptidoglycane et l'incorporation des sous-unités auxiliaires sur le corps du pilus. Afin de polymériser la sous-unité piline majeure en fibre, SrtC reconnaît le motif LPXTG de cette dernière et clive entre la thréonine et la glycine (via sa cystéine nucléophile) (figure 19A). L'intermédiaire acyl-enzyme qui en résulte est résolu par attaque nucléophile du groupe amine de la lysine du motif piline d'une autre piline majeure (figure 19B). Deux pilines majeures sont ainsi pontées covalamment entre elles. Il en va de même pour construire le corps du pilus (figure 19C). L'ancrage du pilus sur le peptidoglycane serait réalisé par une sortase de classe A. En effet, même si cela a été démontré chez *C. diptheriae* (Swaminathan *et al.*, 2007), l'acteur sortasique attachant le pilus à la paroi bactérienne reste encore à confirmer. Le mécanisme par lequel une sous-unité mineure est incorporée sur la structure du pilus demeure à ce jour obscur.

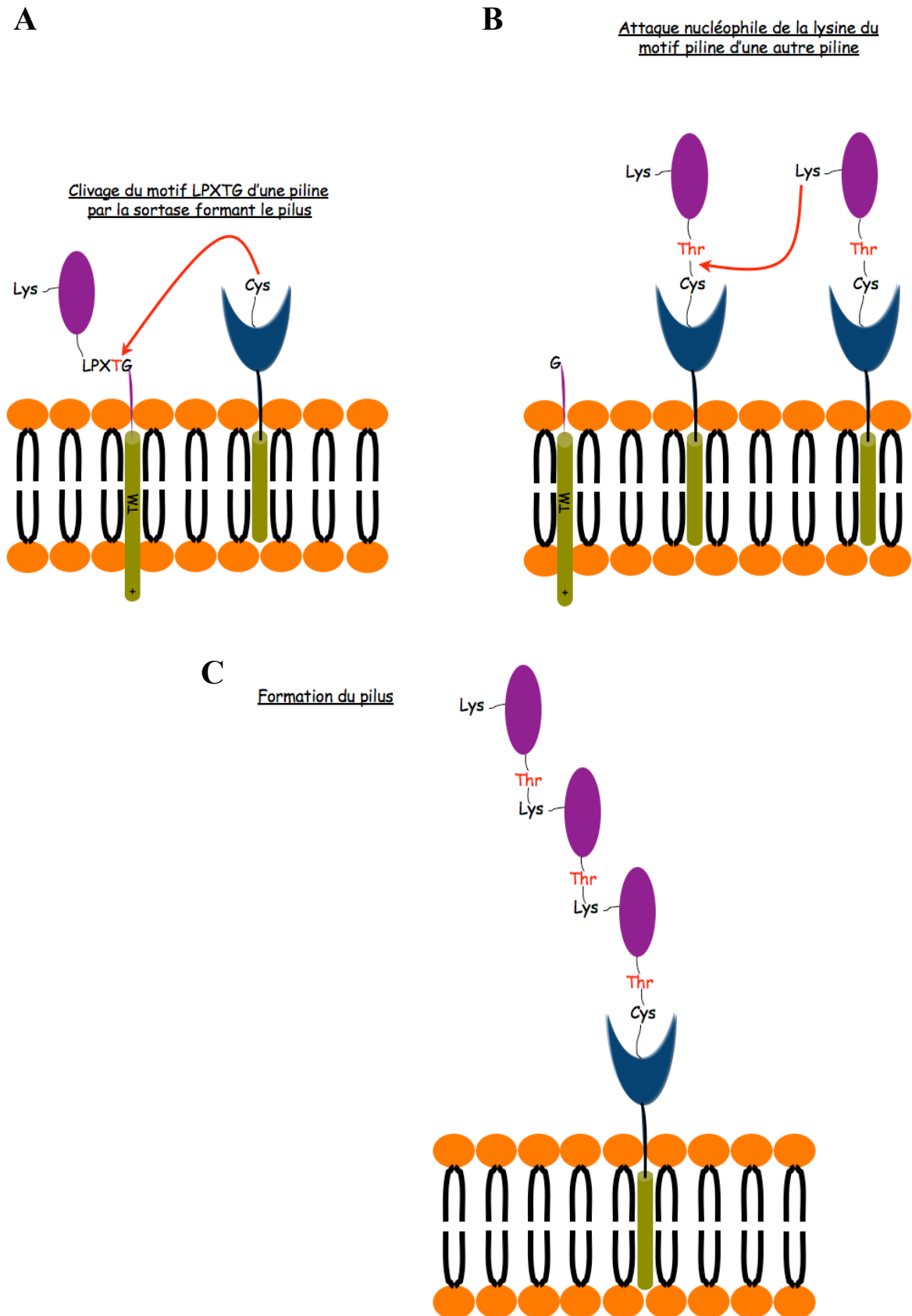


Figure 19 : Biogenèse du pilus chez les bactéries à Gram-positif. A/ La sortase formant le pilus (sortase de classe C) clive le motif LPXTG d'une piline majeure au niveau de la thréonine, générant ainsi un intermédiaire acyl-enzyme. B/ Attaque nucléophile du groupement amine de la chaîne latérale de la lysine du motif piline permettant ainsi de lier covalamment deux pilines entre elles. C/ Formation du pilus.

Curieusement, les sortases de classe C sont plus hétérogènes dans leur séquence que les autres et ne sont présentes que dans certaines souches bactériennes parmi une espèce donnée. Ces observations peuvent suggérer que les régions contenant les gènes des SrtC sont soumises au transfert latéral de gènes et doivent être requises pour une adaptation à une niche spécifique. SrtC-1, SrtC-2 et SrtC-3 de *S. pneumoniae* sur lesquelles nous travaillons font partie de cette classe de sortases.

IV.1.4. Les sortases de classe D (SrtD)

Les sortases de cette classe appartiennent à trois espèces possédant des cycles de différenciations morphologiques: les bacilles, les clostridia et les actinomycetales. Seule SrtD de *B. anthracis* a été caractérisée à ce jour. Le gène codant pour cette sortase est localisé dans un opéron contenant deux protéines régulatrices (SctR et SctS) et une protéine portant le motif LPNTA (BasI). Il a été montré que SrtD n'est exprimée que lors de la sporulation et ne reconnaît que le motif LPNTA de la protéine BasI. SrtD ancre ainsi la protéine BasI sur le peptidoglycane de la cellule mère et sur celui de la spore (Marraffini and Schneewind, 2006). Les sortases de classe D seraient alors impliquées dans l'ancrage de protéines ayant des rôles spécifiques liés à des situations environnementales particulières (comme la sporulation) sur la paroi bactérienne.

IV.2. Informations structurales sur les sortases ancrant des protéines cibles sur la paroi bactérienne

Seules les structures de sortases de classes A et B étaient disponibles au début de ce projet. Nous nous référerons donc, dans cette partie, au repliement et au mécanisme catalytique de ces enzymes.

IV.2.1. Le repliement des sortases de classes A et B

La première structure résolue par cristallographie aux rayons X a été celle de SrtA de *S. aureus* (Zong *et al.*, 2004a; Zong *et al.*, 2004b). SrtA présente un repliement unique,

constitué d'un tonneau central formé de 8 brins β , d'une hélice α et de 3 hélices 3_{10} (figure 20A). Les sortases de classe B dont les structures ont été résolues à ce jour (SrtB de *S. aureus* et *B. anthracis*) adoptent également une structure en tonneau composé de 8 brins β (Zhang *et al.*, 2004), comme le montre la figure ci-dessous.

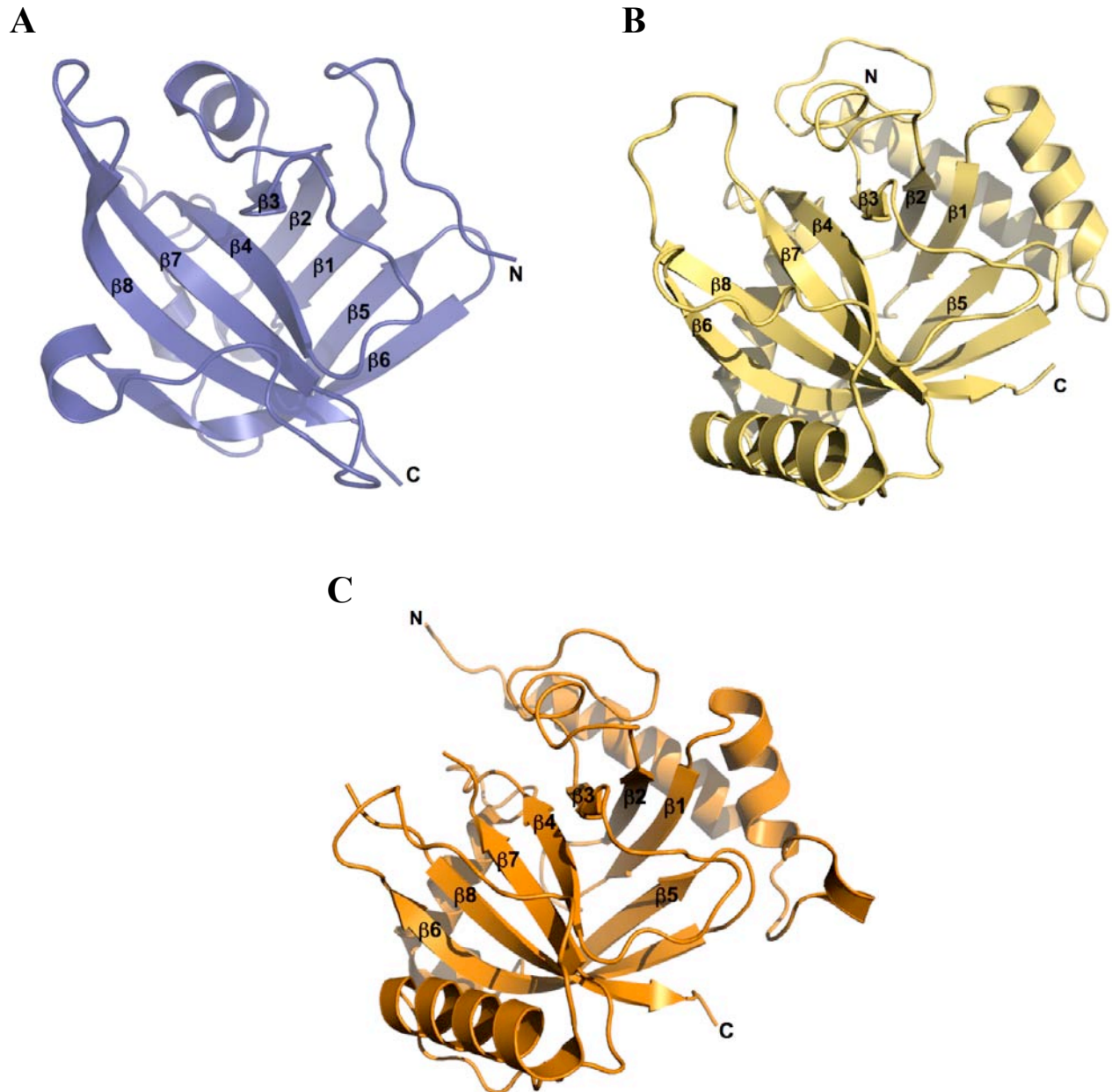


Figure 20 : Structure tridimensionnelle de sortases de classes A et B. A/ Structure de SrtA de *S. aureus* (code PDB : 1T2P). B/ Structure de SrtB de *S. aureus* (code PDB : 1NG5). C/ Structure de SrtB de *B. anthracis* (code PDB : 1RZ2). Ces trois sortases possèdent un repliement unique composé de 8 brins β organisés en un tonneau β .

La présence de calcium dans la structure de SrtA de *S. aureus* a été rapportée. Le calcium est impliqué dans le réarrangement structural d'une boucle désordonnée (boucle entre $\beta 6$ et $\beta 7$), facilitant ainsi la liaison au substrat. L'activité catalytique de SrtA se retrouve ainsi renforcée en présence de cet ion (Naik *et al.*, 2006).

IV.2.2. L'histoire controversée du site actif

Des études biochimiques, réalisées avant la résolution de la structure de SrtA, avaient montré l'importance de la cystéine pour l'activité catalytique ; un remplacement de la cystéine en alanine abolissait, en effet, complètement l'activité enzymatique *in vitro* et *in vivo*, mettant ainsi en exergue la similarité de réaction catalysée par les protéases à cystéines (Ton-That *et al.*, 1999, 2002). L'hypothèse alors émise est que les sortases catalysent le clivage de la séquence peptidique LPXTG avec un mécanisme qui s'apparente à celui de la famille des protéases à cystéines type papaïne. Il est connu que ces dernières possèdent un site actif contenant 3 résidus conservés : une cystéine, une histidine et une asparagine (figure 21). Avant la liaison au substrat, la cystéine est tenue dans un état inactif à travers une interaction ion thiolate/imidazolium avec l'histidine. L'asparagine, par une interaction avec l'histidine, stabilise cette paire d'ion (Vernet *et al.*, 1995).

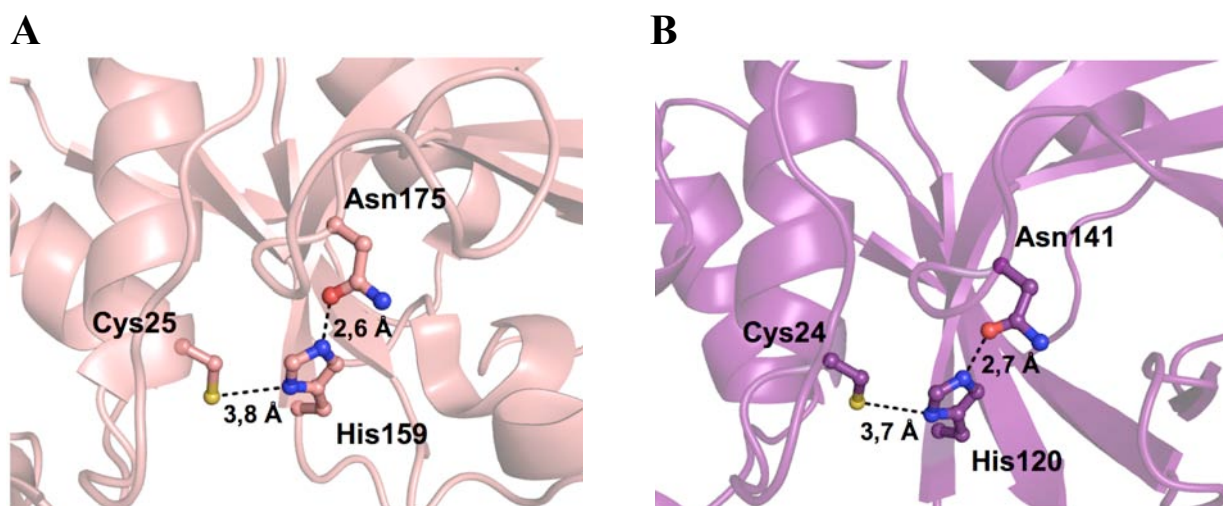


Figure 21 : Site actif des protéases à cystéines. A/ Triade catalytique de la papaïne (code PDB : 1PPN) . B/ Triade catalytique de la staphopapaïne (code PDB : 1CV8).

De ce fait, il a été stipulé que les sortases contenaient probablement une triade catalytique similaire. La recherche de résidus conservés, histidine et asparagine, parmi plusieurs sortases révéla qu'aucun résidu asparagine n'était conservé mais qu'une histidine conservée pourrait faire partie de cette triade. Des études biochimiques sur SrtA de *S. aureus*, dans laquelle l'histidine (His120) a été mutée, ont conforté l'idée que le site actif des sortases contenait une histidine. En effet, cette mutation réduit dramatiquement l'activité catalytique de SrtA *in vivo* et *in vitro* (Ton-That *et al.*, 2002).

La résolution de la structure de SrtA et SrtB en forme apo de *S. aureus* a révélé que la cystéine catalytique et l'histidine conservée occupaient des positions analogues à celles de ces mêmes résidus dans les protéases à cystéine comme la papaïne ou la staphopapaïne (Zhang *et al.*, 2004). Cependant, il s'avéra que l'histidine était trop loin de la cystéine catalytique pour jouer un rôle de paire thiolate/ion imidazolium, et déprotoner le thiol de la cystéine du site actif (figure 22). Des études d'enzymologie, avec la mesure des pKa, ont également conforté le fait que l'histidine ne pouvait former une paire thiolate/ion imidazolium avec la cystéine. Ainsi, l'histidine apparaît comme un acide aminé clé dans l'activité catalytique de la sortase mais son rôle reste à déterminer.

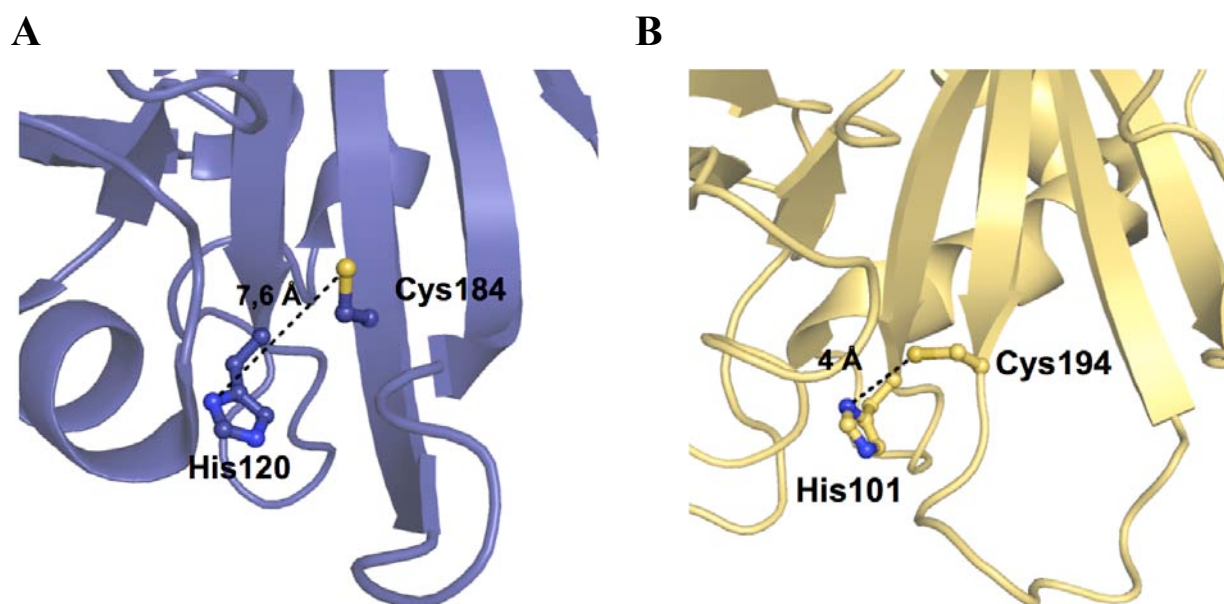


Figure 22 : Sites actifs des sortases A et B de *S. aureus* révélant l'impossibilité de l'histidine de participer à une paire d'ion imadazolium/thiolate. A/ SrtA de *S. aureus*. B/ SrtB de *S. aureus*.

C'est avec la résolution de la structure de SrtA en complexe avec le peptide mimant le motif LPXTG que la compréhension du site actif s'est améliorée. L'existence d'une triade catalytique composée d'une cystéine, d'une histidine et d'une arginine (Cys184, His120 et Arg197 dans SrtA de *S. aureus*) a ainsi pu être mise en évidence (Zong *et al.*, 2004a). Arg197 est en effet le seul résidu conservé ionisable qui pourrait bouger à proximité de Cys184 pour former un couple d'ion catalytique. De plus, cette Arg197 est localisée à proximité de la liaison scissile entre la thréonine et la glycine du peptide (figure 23).

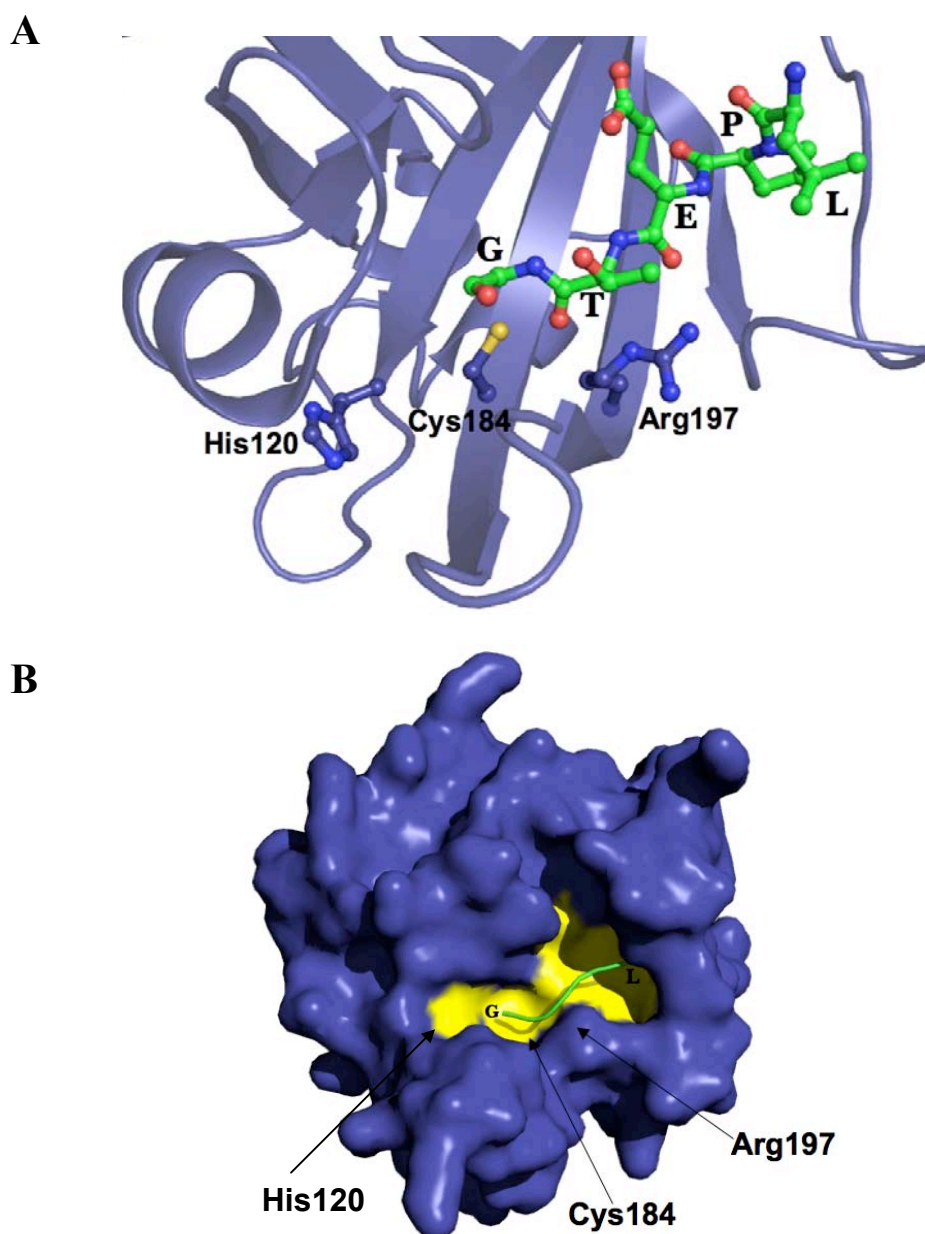


Figure 23 : Structure de SrtA de *S. aureus* complexée avec le peptide LPETG (code PDB : 1T2P). A/ Zoom sur le site actif de SrtA où l'Arg197 se localise près du substrat LPETG. La structure résolue est celle dans laquelle la cystéine catalytique est mutée en alanine, cependant, pour plus de clarté, c'est la cystéine qui est représentée. B/ Représentation de la surface de SrtA ; le peptide LPETG, en vert, se situe dans une poche catalytique, montrée en jaune.

Des études de mutagenèse ont confirmé le rôle important de cet Arg197 puisque le remplacement de cet acide aminé en alanine réduit considérablement l'activité catalytique de SrtA (Marraffini *et al.*, 2004). Par conséquent, il est maintenant admis que le site actif des sortases est composé d'une **triade catalytique Cys-His-Arg** même si le rôle exact de chacun des deux derniers résidus n'est pas totalement compris.

IV.2.3. La triade catalytique

Le site actif, composé de la triade catalytique Cys-His-Arg, se trouve dans une poche allongée plutôt hydrophobe, comme l'indiquent les figures 23 et 24.

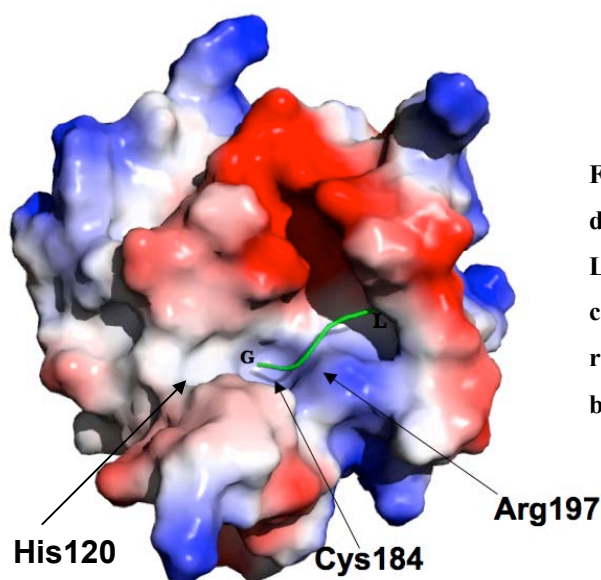


Figure 24 : Représentation de la surface électrostatique de SrtA de *S. aureus* (code PDB : 1T2P). Le peptide LPETG, montré en vert, se localise dans une gorge catalytique plutôt hydrophobe. En rouge sont représentées les zones acides, en bleu, les zones basiques et en blanc, les zones hydrophobes.

Des controverses existent encore aujourd'hui quant au rôle joué par His120 et Arg197. La liaison scissile du peptide, entre la thréonine et la glycine, est située entre les chaînes latérales de Cys184 et Arg197, pointant le fait que Arg197 doit agir comme base pour l'ionisation du groupement thiol de Cys184 durant la catalyse (Zhang *et al.*, 2004 ; Zong *et al.*, 2004a). Une autre hypothèse a cependant ensuite émergé en raison du désaccord existant entre les études de cinétiques et de mutagenèse, qui prédisent un rôle critique pour His120 dans le mécanisme catalytique, et les études structurales qui montre His120 trop éloignée de Cys184 pour former une paire d'ions catalytiques. Arg197 n'agirait ainsi non pas comme une base générale mais stabiliserait l'intermédiaire oxyanion transitoire et/ou serait impliquée dans la liaison et l'orientation du substrat (Frankel *et al.*, 2007 ; Bentley *et al.*, 2008). His120, quant à elle, jouerait un rôle dans la protonation du groupe partant (figure 25).

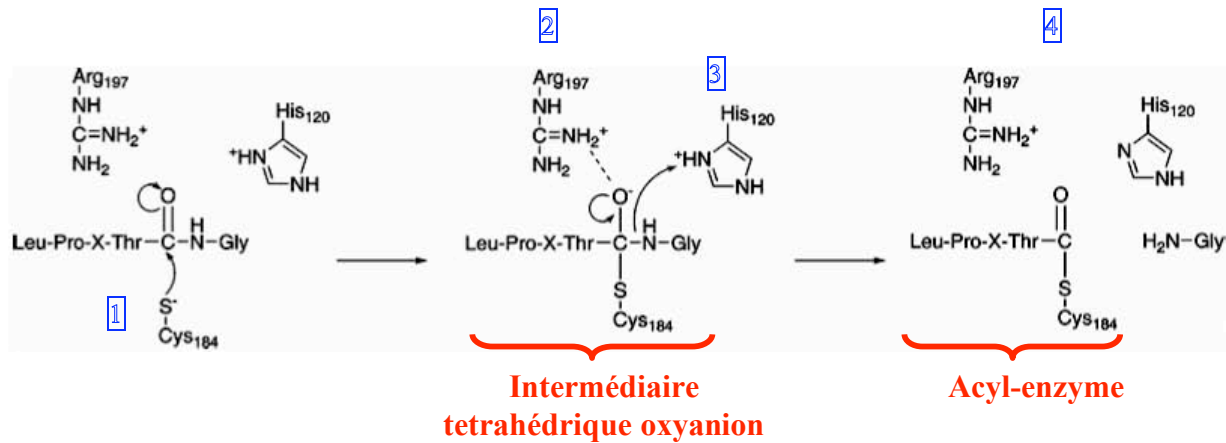


Figure 25 : Modèle suggéré pour la catalyse de SrtA de *S. aureus* (Frankel *et al.*, 2007). Le thiolate de la cystéine nucléophile attaque le groupement carbonyle de la liaison scissile du peptide entre la thréonine et la glycine (1). Un intermédiaire tétrahédrique oxyanion est alors formé et est stabilisé par interaction avec la charge positive de la chaîne latérale de l'arginine (2). La protonation du groupe partant par l'histidine (3) facilite la résolution de l'intermédiaire et la formation de l'acyl-enzyme (4) ;

Contexte et objectif des travaux

Du fait de l'émergence de la résistance aux antibiotiques et de la couverture vaccinale restreinte, des efforts importants de la communauté scientifique sont réalisés afin de découvrir et caractériser de nouvelles cibles thérapeutiques et lutter ainsi contre le pneumocoque. Les facteurs de virulence de celui-ci constituent des cibles attractives pour le développement de nouveaux traitements. La découverte des pili comme étant des acteurs impliqués dans sa pathogénie incite à étudier ce système plus en profondeur.

L'objectif de la thèse consistait en la caractérisation structurale mais aussi fonctionnelle des divers composants du pilus de *Streptococcus pneumoniae*. Anne-Marie Di-Guilmi, qui a initié le projet, aidée de Lamy El Mortaji du Laboratoire d'Ingénierie Macromoléculaire, ont réalisé les constructions de toutes les protéines du pilus (RrgA, RrgB, RrgC, SrtC-1, SrtC-2 et SrtC-3) avant mon arrivée sur le projet. J'ai rejoint le projet « pilus » 6 mois après mon arrivée au Laboratoire des Protéines Membranaires du fait d'un remaniement de mon sujet de thèse initial, suite à des publications de nos concurrents dans des journaux de renommée internationale. Peu de choses étaient connues sur le pilus de *S. pneumoniae* lors de notre début d'étude. Seules des études de microbiologie génétique, révélant la présence de pili à la surface pneumococcique et la composition de celui-ci, étaient disponibles. De plus, aucune information structurale sur les différentes protéines de l'îlot *rlrA* existait. Notre objectif, s'inscrivant dans une logique de caractérisation de nouvelles cibles thérapeutiques, était d'une part, d'étudier structuralement les diverses protéines du pilus et, d'autre part, d'étudier la formation du pilus par les sortases.

Une piste prometteuse concernant la structure de SrtC-1 était disponible et un jeu de donnée cristallographique natif avait été collecté lors de mon arrivée sur le projet. Mon premier objectif de thèse consistait à travailler sur SrtC-1 et à résoudre sa structure par cristallographie aux rayons X. Parallèlement à la résolution de sa structure, nous nous sommes intéressés à sa fonction dans la biogenèse du corps du pilus, ainsi qu'à celle des autres sortases. En effet, déterminer le ou les acteurs permettant la polymérisation du pilus ouvrait la voie dans la caractérisation de cibles thérapeutiques potentielles. C'est vers une approche biochimique avec l'utilisation des protéines recombinantes du pilus (approche originale du fait qu'aucune étude de ce genre n'est encore utilisée pour étudier la formation de pili) que nous nous sommes tournés pour étudier le rôle des trois sortases de l'îlot *rlrA* de *S. pneumoniae*. L'ensemble de ces résultats, ont été décrits dans un article publié dans *Structure* (Déc 2008).

La deuxième partie de ce travail consistait à mieux caractériser SrtC-1 afin d'améliorer nos connaissances sur le fonctionnement de sortases impliquées dans la formation de pili. Ces analyses ont été décrites dans un second article qui est actuellement en préparation.

Matériel et Méthodes

I. Biologie moléculaire

I.1. Clonage

Le clonage de RrgA, RrgB, RrgC, SrtC-1, SrtC-2 et SrtC-3 a été réalisé par Anne-Marie Di-Guilmi au Laboratoire d'Ingénierie des Macromolécules à l'IBS. L'ADN génomique de la souche de *S. pneumoniae* encapsulé TIGR4 a été utilisé comme matrice d'amplification au cours d'une réaction de PCR (« Polymerase Chain Reaction »). Les régions génomiques contenues dans l'îlot de pathogénicité *rlrA* et correspondant aux protéines RrgA (acides aminés 38-869), RrgB (acides aminés 30-633), RrgC (acides aminés 21-368), SrtC-1 (acides aminés 17-228), SrtC-2 (acides aminés 9-220), SrtC-3 (acides aminés 32-254) ont été ainsi amplifiées puis clonées dans le vecteur d'expression pLim (Protein Expert, Grenoble). Ce plasmide, portant la résistance à l'ampicilline, possède une séquence codante pour une étiquette hexahistidine en N-terminal clivable par la protéase Tev. Des cellules *E. coli* BL21(DE3)-RIL (Invitrogen), portant la résistance au chloramphénicol, ont été ensuite transformées avec les différents vecteurs pour l'expression des protéines.

I.2. Mutagenèse dirigée

I.2.1. Construction des mutants *SrtC-1C193A*, *SrtC-1H131D*, *SrtC-R202E*, *SrtC-1D58GW60G*, *SrtC-1D58GW60GC193A*

Les mutants de SrtC-1 ont été construits en utilisant le kit de mutagenèse dirigée « Quick Change Mutagenesis kit » (Stratagène). Le plasmide pLim double brin contenant le gène d'intérêt a été utilisé comme matrice d'ADN parentale. 40 ng d'ADN parental, 125 ng de chaque amorce, 10 mM dNTP, 2,5 unités de polymérase Pfu Turbo et le tampon de la polymérase sont utilisés pour la réaction. Une PCR est effectuée pour introduire la mutation. Une digestion subséquente de 2 heures à 37°C avec l'enzyme DpnI, spécifique de l'ADN méthylé et semi méthylé, élimine l'ADN parental et sélectionne l'ADN contenant la mutation. Après transformation dans des cellules compétentes d'*E. coli* (Top10 ou NEB5 α , (BioLabs)), l'ADN plasmidique est extrait grâce au kit « Nucleospin Plasmid » (MACHEREY NAGEL)

et séquencés par Cogenics afin de vérifier la présence de la mutation. Les diverses amorces nucléotidiques utilisées ainsi que les cycles de PCR effectués sont détaillés en annexe A. Il est à noter que pour réaliser le triple mutant SrtC-1D58GW60GC193A, nous avons utilisé, comme matrice d'ADN, le plasmide pLim/SrtC-1.

1.2.2. Construction des mutants SrtC-1_{lid2} et SrtC-2_{lid1}

Notre but ici était d'inverser les couvercles de SrtC-1 et SrtC-2. Nous appelons mutant SrtC-1_{lid2}, la sortase SrtC-1 dont le couvercle (acides aminés 52 à 78) a été remplacé par celui qui possède SrtC-2. À l'inverse, Nous appelons mutant SrtC-2_{lid1}, la sortase SrtC-2 dont le couvercle (acides aminés 44 à 70) a été remplacé par celui de la SrtC-1. Les amorces nucléotidiques ainsi que les cycles PCR utilisées sont décrits en annexe A.

Dans un premier temps, une PCR a été effectuée avec des amorces oligonucléotidiques pour amplifier la région génomique correspondant aux couvercles et contenant les régions flanquantes du gène cible opposé. Cette première étape a servi à générer les amorces pour la seconde PCR. Le plasmide pLim double brin contenant le gène d'intérêt (SrtC-1 ou SrtC-2) est utilisé comme matrice d'ADN parentale. La réaction contient 5 ng d'ADN parental, 20 ng de chaque amorce, 10 mM dNTP, 2,5 unités de polymérase Pfu Turbo et le tampon de la polymérase. Une étape de purification du produit PCR est ensuite effectuée. Pour ce faire, les produits PCR sont chargés sur gel agarose, coloré au BET pendant 5 minutes puis les bandes correspondantes sont coupées. Le kit « Qiaquick Gel » (Qiagen) est utilisé pour purifier les produits PCR. Dans un second temps, l'étape de mutagenèse dirigée est réalisée grâce au kit « Quick Change XL Site Directed Mutagenesis » (Stratagène). Le plasmide pLim double brin contenant le gène d'intérêt (SrtC-1 ou SrtC-2) est utilisé comme matrice d'ADN parentale. La réaction contient 100 ng d'ADN parental, 120 ng d'amorces précédemment réalisées (produits PCR), 10 mM dNTP, 2,5 unités de polymérase Pfu Turbo et le tampon de la polymérase. Après PCR, une digestion de 3 heures à 37°C avec l'enzyme DpnI est par la suite effectuée. Après transformation dans des cellules compétentes *E. coli* NEB5α (BioLabs), l'ADN plasmidique est extrait grâce au kit « Nucleospin Plasmid » (MACHEREY NAGEL). La présence de l'insert est vérifiée par contrôle de la taille de ce dernier en réalisant une digestion avec les enzymes de restriction, PstI et NcoI, qui correspondent au site de clivage du gène d'intérêt. Les plasmides positifs sont alors séquencés par Cogenics afin de vérifier la présence des séquences attendues.

II. Biochimie

II.1. Surexpression de la protéine d'intérêt

Les protéines RrgA, RrgB, RrgC, SrtC-1, SrtC-2, SrtC-3 ainsi que l'ensemble des mutants ont été exprimées et purifiées selon le protocole suivant.

Une culture de cellules BL21(DE3)-RIL en milieu TB (Terrific Broth), supplémenté de 100 µg/mL d'ampicilline (Euromedex) et de 34 µg/mL de chloramphénicol (Euromedex), est préparée en diluant une préculture à saturation au 1/20^{ème}. Les cellules sont laissées à 37°C, sous agitation (200 rpm), jusqu'à ce qu'elles atteignent une DO_{600nm}=0,6 U.A. L'expression des protéines est alors induite par l'ajout de 1mM d'IPTG (Euromedex). Après 3 heures d'agitation et d'induction à 37°C, les cellules sont centrifugées à 5500 rpm pendant 12 minutes. Le culot bactérien est repris dans le tampon de lyse (50 mM Tris-HCl pH 8,0, 200 mM NaCl, 20 mM Imidazole). Après ajout de 1 mM de PMSF, 1 µM de pepstatine et 0,2 µM d'aprotinine, les cellules sont lysées par sonication (2 minutes de sonication totales par pulses de 2 secondes espacés de 10 secondes). Le lysat est ensuite centrifugé à 18000 rpm pendant 30 minutes afin de séparer les protéines solubles du matériel insoluble.

II.2. Purification de la protéine d'intérêt

II.2.1. Chromatographie d'affinité sur colonne HisTrapTMHP

Le surnageant, contenant les protéines solubles, est déposé sur une colonne d'affinité HisTrapTMHP (Amersham Biosciences) préalablement équilibrée avec le tampon de lyse. Après un lavage équivalant à 20 volumes de la colonne avec le tampon de lyse, les protéines sont éluées par un gradient linéaire de 20 à 500 mM Imidazole, sur 20 volumes de colonne. Les fractions sont ensuite analysées sur gel SDS.

II.2.2. Coupure de l'étiquette hexahistidine

L'étiquette histidine est clivable grâce à la présence du site de reconnaissance de la protéase Tev (ENLYFQG) situé entre le His₆ et la protéine d'intérêt.

Les fractions contenant la protéine d'intérêt sont incubées avec la Tev (rapport massique Tev/protéine de 1 :10) et dialysées sur la nuit à 4°C dans le tampon de lyse. L'activité de la protéase est ensuite inhibée par l'ajout de 1 mM de PMSF. L'échantillon clivé est de nouveau chargé sur colonne HisTrapTMHP afin de ne récupérer que la protéine non étiquetée et d'éliminer la Tev (en effet, celle-ci, possédant une étiquette histidine, se fixera sur la colonne).

II.2.3. Chromatographie d'exclusion de taille

Les fractions contenant la protéine d'intérêt non étiquetée sont chargées sur une colonne d'exclusion de taille, préalablement équilibrée dans le tampon 50 mM Hepes-HCl pH 7,5, 150 mM NaCl. Suivant la quantité de protéine obtenue, une colonne d'exclusion Superdex 200 HR 10/30 (Amersham Biosciences) ou une colonne d'exclusion HiLoad 16/60 Superdex 200 (Amersham Biosciences) est utilisée.

II.2.4. Production de SrtC-1 et RrgB sélénées

La même souche bactérienne BL21(DE3)-RIL est utilisée pour la production des protéines SrtC-1 et RrgB sélénées. Les cellules de la préculture sont dans un premier temps centrifugées pendant 15 minutes à 3500 rpm, lavées dans du milieu minimum M9 (Annexe B), puis utilisées pour inoculer 400 mL de milieu minimum M9 au 1/5^{ème}. La culture est laissée à 37°C sous agitation jusqu'à atteindre une DO_{600nm}=0,7 U.A. Cette culture est ensuite utilisée pour inoculer, au 1/5^{ème}, 2L de milieu minimum M9, supplémenté de 20 mL de solution d'acides aminés 50x (Annexe B). À une DO_{600nm}=0,6 U.A, 120 mg des acides aminés L-Lysine, L-Phénylalanine, L-Thréonine, 60 mg des acides aminés L-Isoleucine, L-Leucine, L-Valine et 70 mg de Sélénométhionine sont ajoutés. L'incubation à 37°C et sous agitation se poursuit alors durant 25 minutes. L'induction de l'expression des protéines se réalise ensuite par l'ajout de 1 mM IPTG pendant une nuit à 16°C, sous agitation.

II.3. Formation *in vitro* des fibres de RrgB

Les protéines purifiées RrgB et SrtC-1 sont mélangées à un ratio massique respectif de 1:2 et incubées à 37°C pendant une nuit dans un thermobloc. Un protocole similaire est utilisé avec SrtC-2, SrtC-3 et les mutants de SrtC-1. Un volume constant est à chaque fois effectué. Après incubation, les échantillons sont chauffés à 100°C pendant 10 minutes en présence de bleu de Coomassie et de 1% SDS. La présence ou non de fibres de RrgB est analysée par Western Blot. Pour ce faire, après migration des échantillons sur un gel de gradient 4-12% SDS (gel fabriqué au laboratoire ou gel commercial (BioRad)), un transfert sur membrane (Transfer-Blot^R Medium, Bio-Rad) est réalisé pendant 1h10 à 300 mA. La membrane est ensuite incubée pendant 1h dans du tampon PBS Tween 0,3% additionné de 5% de lait en poudre (Régilait). Les anticorps polyclonaux de souris (anti-RrgB, anti-SrtC-1, anti-SrtC-2 ou anti-SrtC-3) sont subséquemment ajoutés à une dilution de 1/5000. Après 1h30 d'incubation, 3 lavages successifs de 10 minutes chacun dans du PBS Tween 0,3% sont réalisés. La membrane est ensuite incubée avec des anticorps secondaires anti-souris couplés à la HRP (Sigma, 1/10000) pendant 1 nuit à 4°C. Après 3 lavages successifs de 10 minutes dans du PBS Tween 0,3%, la membrane est révélée par chimiluminescence, soit directement sur la membrane à l'aide de tablette SIGMA-FASTTM 3,3'-Diaminobenzidine (SIGMA), soit avec le réactif ECL (Bio-Rad).

III. Analyses biophysiques

III.1. Spectrométrie de masse

Les analyses de spectrométrie de masse ont été réalisées au Laboratoire de Spectrométrie de Masse des Protéines à l'IBS afin de vérifier le poids moléculaire des mutants et de mettre en évidence le complexe SrtC-1/RrgB. C'est une méthode d'analyse permettant de mesurer le rapport masse/charge (m/z) de particules ionisées en phase gazeuse, et ainsi d'obtenir le poids moléculaire de protéines ou de peptides. Les sources d'ionisations MALDI (Matrix-Assisted Laser Desorption) et ESI (ElectroSpray Ionisation), ainsi qu'un analyseur TOF (Time Of Flight) ont été utilisées dans notre cas. Ce dernier mesure le temps que met un ion, soumis à une tension préalable, à parcourir une distance donnée. La séparation des ions de ce type d'analyseur est fondée sur leur différence de vitesse. À la sortie de la source d'ionisation, tous les ions possèdent la même énergie cinétique, par conséquent, les ions les plus légers ont une vitesse plus grande que les ions les plus lourds. Le temps nécessaire pour que l'ion atteigne le détecteur est ainsi uniquement dépendant du rapport m/z .

L'analyse de la nature des fibres a été effectuée par nanoLC-MS/MS au Laboratoire d'Etude de la Dynamique des Protéomes au CEA, grâce à la plate-forme « EDyP-Service » disponible au Laboratoire d'Etude de la Dynamique des Protéomes du CEA Grenoble. Cette technique permet l'analyse d'échantillons biologiques à partir d'un gel d'électrophorèse. Après découpe des bandes protéiques souhaitant être analysées, les protéines sont soumises à une digestion enzymatique. Les peptides issus de cette hydrolyse enzymatique vont ensuite être séparés par chromatographie liquide en phase inverse suivant leur hydrophobicité. Après avoir été ainsi séparés, les peptides sont introduits séquentiellement dans le spectromètre de masse afin d'être ionisés et fragmentés. L'analyse du rapport m/z de chaque peptide fragmenté permet d'accéder à la séquence des peptides issus de la digestion trypsique et ainsi de caractériser les bandes protéiques découpées sur gel.

III.2. Séquençage protéique N-terminal

Le séquençage protéique N-terminal, réalisé au Laboratoire d'Enzymologie Moléculaire de l'IBS, est effectué selon le principe de la dégradation d'Edman grâce à un séquenceur en phase gazeuse modèle 492 (Applied Biosystems). Une électrophorèse en gel de polyacrylamide en condition dénaturante est réalisée avec 10 µg de protéine purifiée. Le gel est ensuite lavé dans 10 mM CAPS pH 11.0 (Sigma), 10 % méthanol. La membrane PVDF Immun-BlotTM (Bio-Rad), préalablement lavée dans du méthanol, est imbibée dans ce même tampon et déposée sur le gel de polyacrylamide. Les protéines sont alors transférées du gel vers la membrane après application d'un courant de 300 mA pendant 1 heure à l'aide d'un appareil à transfert (Bio-Rad). La membrane est ensuite colorée dans du bleu de Coomassie dans 50 % méthanol, 1 % acide acétique pendant 5 minutes puis décolorée dans 50 % méthanol (2 lavages de 10 minutes chacun). Les bandes souhaitant être analysées sont découpées et soumises au séquençage N-terminal. Ce type d'analyse a permis de confirmer l'identité de SrtC-1 et de RrgB, notamment pour les deux bandes présentes sur gel SDS et qui s'avèrent être toutes deux RrgB.

III.3. Étude de la thermostabilité des protéines par fluorescence : Thermal Shift Assay (TSA)

III.3.1. Principe

Le Thermal Shift Assay est une technique qui permet de mesurer la dénaturation des protéines induite par un gradient de températures. C'est grâce à des colorants fluorescents sensibles à leur environnement, comme le Sypro Orange (Molecular Probes), que la dénaturation d'une structure protéique peut être suivie. En effet, lorsqu'un tel colorant est ajouté à une protéine native en solution, il est exposé à un environnement aqueux et l'émission de fluorescence est faible. En revanche, lorsque la température augmente, la protéine se déstructure et expose ses régions hydrophobes ; le colorant se lie à ces régions hydrophobes, désormais accessibles, et devient alors fortement fluorescent. Cela se traduit donc par une augmentation du signal de fluorescence du colorant indiquant une dénaturation de la protéine.

La courbe de fluorescence typique obtenue au cours de la cinétique de dénaturation thermique d'une protéine est représentée figure 26. La température de transition thermique, appelée T_m , correspond au point d'inflexion de la courbe de fluorescence. Afin de la déterminer plus aisément, la dérivée première des données brutes est calculée. La T_m correspond alors au maximum de la courbe obtenue. Les variations de la valeur de la T_m sont indicatives d'un changement dans la stabilité de la protéine. Un déplacement de T_m vers les températures élevées indique une stabilisation de la protéine par augmentation de sa structuration, concomitante à une réduction de sa flexibilité conformationnelle. En revanche, une diminution de la T_m caractérise une déstabilisation conformationnelle de la protéine. (Pantoliano *et al.*, 2001 ; Geerlof *et al.*, 2006).

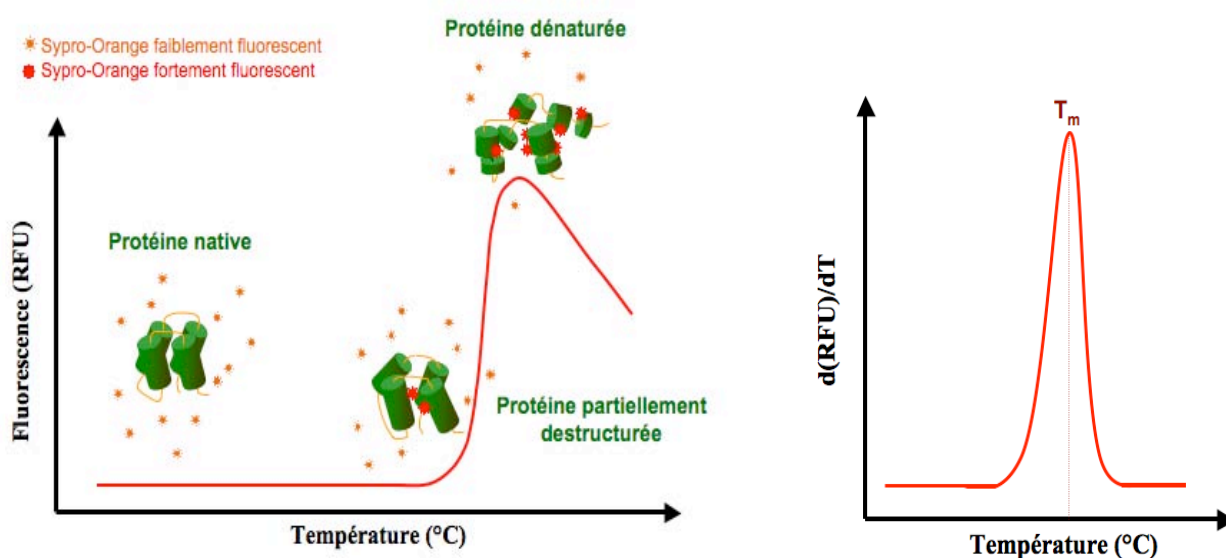


Figure 26 : Principe du Thermal Shift Assay. La fluorescence est exprimée en unité arbitraire (RFU). Le calcul des dérivées premières de chaque point de la courbe de fluorescence (à gauche) permet de déterminer plus facilement la T_m . Ceci est représenté à droite.

III.3.2. Acquisition des données

Le TSA a été utilisé pour étudier l'influence des mutations sur la thermostabilité des protéines. Il a été également utilisé pour analyser l'effet des peptides sur SrtC-1 et ses mutants. Les expériences de TSA ont été réalisées sur un appareil à PCR en temps réel Thermocycler iQ5 (Bio-Rad) disponible au Laboratoire d'Ingénierie des Macromolécules de

l'IBS. Le colorant utilisé est le Sypro Orange. Les tests sont effectués dans des microplaques 96 puits sur un volume de 25 μ l (23 μ l de protéines à 5 mg/ml minimum + 2 μ l de Sypro Orange 50X). Les cinétiques de dénaturation sont réalisées de 20 °C à 100 °C, avec un pas de 1°C par minute sauf pour le triple mutant SrtC-1CD58GW60G193A où les échantillons sont chauffés de 20°C à 79°C avec un pas de 0,5°C par minute. L'excitation du Sypro Orange est faite à 470 nm alors que son émission est mesurée à 570 nm. Les données brutes (la courbe de fluorescence) sont enregistrées et un programme permet ensuite de calculer la dérivée première de chaque point de cette courbe.

III.4. Microscopie électronique

III.4.1. Principe

Les études de microscopie électronique ont été réalisées au Laboratoire de Microscopie Electronique et Structurale à l'IBS, en collaboration avec l'équipe de Guy Schoehn. C'est une technique utilisant des électrons comme rayonnement, et non des photons comme c'est le cas pour la microscopie optique. Ce faisceau d'électrons est focalisé sur l'échantillon par un système de lentilles magnétiques. Il interagit alors avec l'échantillon suivant l'épaisseur, la densité ou la nature chimique de celui-ci. L'image obtenue est ensuite transformée en image photonique. Tout comme en microscopie optique où des colorants sont utilisés pour augmenter le contraste des images d'échantillons, en microscopie électronique, ce sont des métaux lourds qui sont utilisés. Le sel d'atome lourd va se fixer préférentiellement à la surface des particules biologiques adsorbées. Un fort contraste est ainsi généré entre les zones riches en métaux lourds (qui vont dévier les électrons) et les autres régions ; l'échantillon biologique apparaîtra alors plus clair que ce qui l'entoure. Cette technique, appelée coloration négative, a été utilisée pour visualiser les fibres de RrgB formées *in vitro*.

III.4.2. Analyse des fibres de RrgB

Après une nuit d'incubation à 37°C, le mélange SrtC-1+RrgB est chargé sur une colonne d'exclusion de taille Superdex 200 (Amersham Biosciences). Les fractions sont ensuite analysées par Western Blot avec les anticorps anti-RrgB ; celles correspondant aux bandes de haut poids moléculaires sont soumises à l'analyse de microscopie électronique.

4 μ l de l'échantillon protéique, à une concentration approximative de 0,05 mg/ml, sont adsorbés à l'interface carbone/mica puis colorés négativement avec 2% de silicotungstate de sodium. La surface carbone/échantillon, débarrassée de la surface de mica, est ensuite déposée sur une grille de cuivre. Les micrographes sont enregistrés à faible dose avec un microscope Phillips CM12 Lab6 à 120kV et avec un grossissement de 45 000 x. Ils sont numérisés en utilisant un scanneur Zeiss avec une taille de pixel de 7 nm.

Les masses des fibres de RrgB peuvent être évaluées à partir des clichés de microscopie, où 1 pixel = 1,17 Å. Si on considère la fibre comme étant un cylindre, le volume de celle-ci correspond à la longueur multipliée par la surface de section (πr^2 , où r représente la moitié du diamètre de la fibre). Par exemple, pour la fibre représentée sur la figure 36, image d, sa longueur mesurée est de 120 Å et sa hauteur de 42 Å. Son volume est donc de $120 \times \pi \times 21^2$ soit d'approximativement 166 253 Å³. Si on suppose que la densité moyenne d'une protéine est de 0,86 Da/ Å³, cette fibre a une masse d'environ 143 kDa

IV. Études cristallographiques des protéines SrtC-1, SrtC-3 et RrgB

IV.1. Introduction à la cristallographie des rayons X

Pour observer l'arrangement des atomes constituant les macromolécules que nous voulons étudier, il est nécessaire d'utiliser un rayonnement dont la longueur d'onde est de l'ordre de grandeur de l'objet à observer. Les distances inter-atomiques dans une molécule étant de l'ordre de l'Angström (10^{-10} m), les rayons X, ondes électromagnétiques de longueur d'onde de l'ordre de l'Angström, sont donc propices à l'étude des protéines par cristallographie. L'interaction entre ces rayons et une molécule génère cependant une diffusion de trop faible intensité pour être mesurée. Il s'avère donc nécessaire d'utiliser un cristal de protéines, dans lequel ces dernières sont empilées de manière ordonnée, qui va permettre aux ondes diffractées d'être cohérentes et à la somme de leurs intensités d'être ainsi observables.

IV.1.1. Cristallogenèse

La cristallisation est un processus d'ordination qui implique un rapprochement des macromolécules en augmentant les interactions entre ces dernières et en excluant une partie du solvant. La méthode utilisée est celle de la diffusion de vapeur en goutte suspendue (McPherson *et al.*, 1995) dans laquelle une goutte formée du mélange de la protéine avec une solution de cristallisation (contenant des agents précipitants et/ou solubilisants) est placée sur un bouchon de plastique. Ce dernier est ensuite scellé au-dessus d'un réservoir contenant un grand volume de la solution de cristallisation (figure 27A). L'équilibre des concentrations s'établit entre la goutte et le réservoir par diffusion des espèces volatiles entraînant ainsi une diminution du volume de la goutte. En résulte donc une augmentation de la concentration de la protéine et de l'agent précipitant dans la goutte.

Le diagramme de phase d'une protéine, présenté dans la figure 27B, montre que lors des étapes de concentration de la protéine et de l'agent précipitant, la solubilité de la protéine diminue jusqu'à atteindre une zone de sursaturation favorable à l'obtention de cristaux.

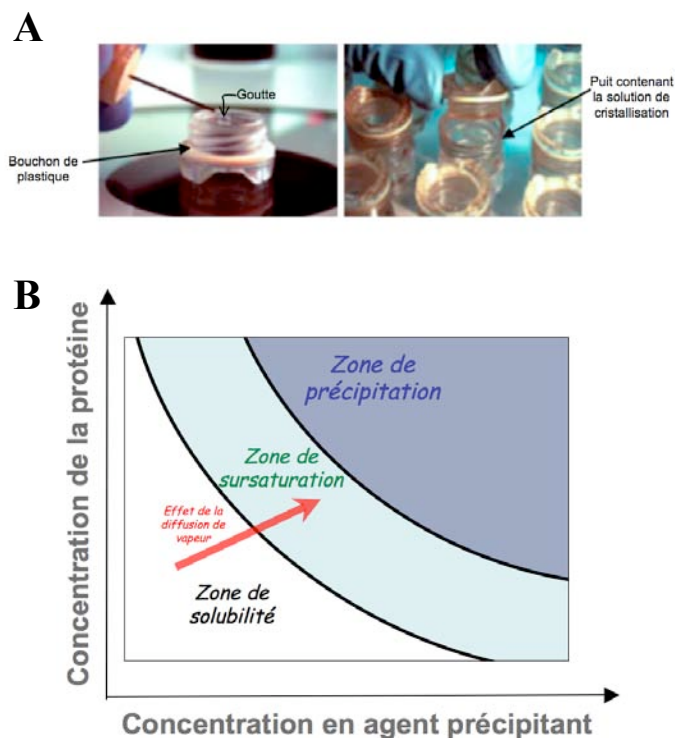


Figure 27 : A/ Mise en oeuvre expérimentale de la technique de diffusion de vapeur par goutte suspendue. B/ Diagramme de phase décrivant le processus de cristallisation d'une protéine par diffusion de vapeur

Les cristaux formés sont soumis à un rayonnement X provenant d'un générateur de laboratoire ou d'un synchrotron. Cependant, avant de soumettre les cristaux aux rayons X, ceux-ci sont refroidis dans l'azote liquide afin de limiter les effets dus aux radiations. En effet, le faisceau de rayons X provoque l'échauffement local du cristal et engendre la formation de radicaux libres qui, à leur tour, endommagent les molécules du cristal. La propagation des radicaux libres dans le cristal est ainsi ralentie par le refroidissement de ce dernier dans l'azote liquide. Préalablement à leur refroidissement dans l'azote liquide, les cristaux sont plongés dans une solution dite cryoprotectrice afin de prévenir la formation de cristaux de glace qui les dégraderaient.

IV.1.2. Le principe de diffraction des rayons X par un cristal de protéine

IV.1.2.1. Note préliminaire sur les cristaux de protéines

Un cristal est la répétition périodique d'un motif géométrique dans les trois directions de l'espace (figure 28 A et B). Le plus petit volume translaté dans ces directions, qui permet de générer un cristal entier, est appelé maille élémentaire. Elle est définie par trois vecteurs \vec{a} , \vec{b} et \vec{c} non colinéaires et trois angles notés α , β , γ , (figure 28C). L'ensemble des translations

permettant de reconstituer le cristal à partir de la maille constitue le réseau. La périodicité de la structure d'un cristal est donc représentée par un ensemble de points régulièrement disposés. Cet ensemble est appelé réseau cristallin et les points le constituant sont appelés nœuds du réseau.

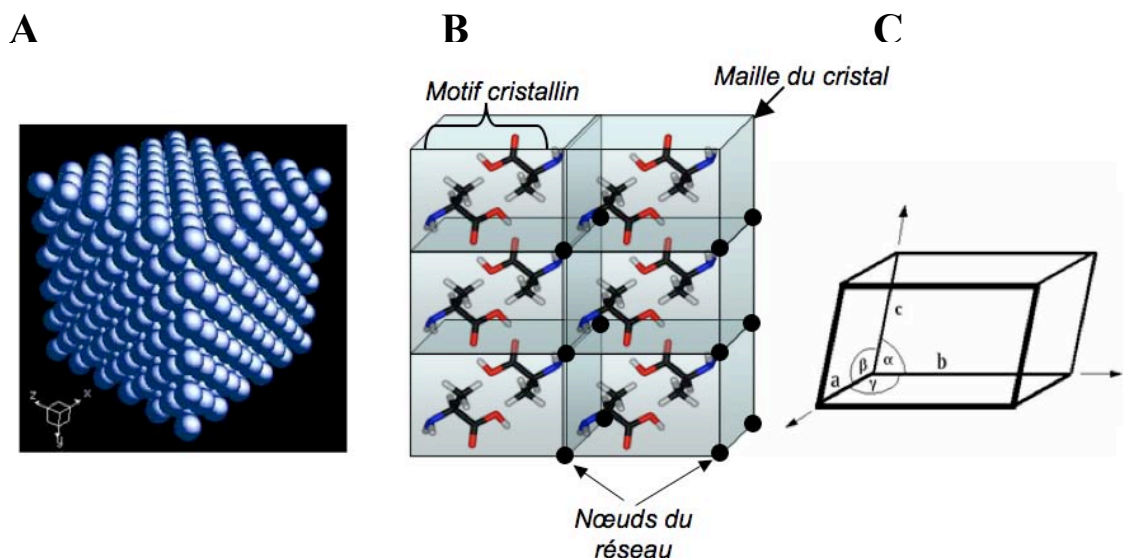


Figure 28 : A/ Modèle d'un cristal à trois dimensions. B/ Représentation mathématique d'un cristal. C/ Maille élémentaire et paramètres à trois dimensions.

Il existe 14 types différents de maille élémentaires pour décrire l'arrangement de points en 3 dimensions qui créent 14 réseaux appelés réseaux de Bravais. Des opérations de symétries (rotation au tour d'un axe, inversion et réflexion) peuvent être appliquées ensuite dans la maille. Elles appartiennent à un des 32 groupes symétriques qui combinés aux 14 réseaux de Bravais forment les 230 groupes d'espace qui représentent toutes les combinaisons d'opérations de symétrie applicables au motif cristallin. Les protéines étant des molécules chirales, les opérations de symétrie de type réflexion et inversion ne peuvent exister dans leurs cristaux, ce qui réduit à 65 le nombre de groupes d'espace possibles en cristallographie des protéines. On définit ainsi l'unité asymétrique qui est le plus petit volume permettant de reconstituer la maille par symétrie.

IV.1.2.2. Diffraction des rayons X par les protéines du cristal

L'interaction entre le rayonnement X et les électrons des protéines dans le cristal se traduit par la diffusion de ce rayonnement dans certaines directions privilégiées. Ceci est

appelé le phénomène de diffraction. Du fait de l'organisation régulière du cristal, dans certains endroits de l'espace, les ondes s'annulent (interférences destructives), et dans d'autres, les ondes s'additionnent et l'on a une intensité positive. L'enregistrement des faisceaux diffractés permet d'obtenir un diagramme de diffraction qui permet, par une interprétation mathématique, de déterminer la symétrie du cristal, les paramètres de maille et la structure du cristal, c'est à dire connaître la position et la nature de tous les atomes du motif.

Les bases mathématiques qui caractérisent le processus de diffraction ont été formulées principalement par Laue, Bragg et Eswald au début du XX^{ème} siècle. Bragg a montré que le phénomène de diffraction observé dans un cristal pouvait être considéré comme la réflexion du faisceau incident par une série de plans parallèles (les plans réticulaires du cristal), comme le ferait une série de miroirs. Une infinité de plans réticulaires, parallèles, contient l'ensemble des nœuds du réseau du réseau cristallin. Chaque ensemble de ces plans est identifié par les indices de Miller hkl . Les rayons X réfléchis par des plans adjacents parcourent des distances différentes et Bragg a montré qu'un faisceau primaire de longueur d'onde λ , frappant l'ensemble de plans parallèles d'indice hkl et de distance d'espacement d_{hkl} avec un angle θ , sera réfléchi et repartira selon le même angle θ (figure 29) seulement si la condition géométrique suivante est rencontrée :

$$2d \sin \theta = n\lambda,$$

où (n) est un entier. Ainsi, si les rayons X frappent les plans parallèles avec un angle qui ne satisfait pas cette équation, les rayons adjacents seront diffractés avec des phases différentes donnant lieu à des interférences destructrices. Seulement lorsque cette équation est satisfaite, des interférences constructives se produisent et s'additionnent pour être observées sur le cliché de diffraction sous forme de tache de diffraction ou réflexion. La position de chaque tache se rapporte à une famille de plans réticulaires et est définie par ses indices de Miller.

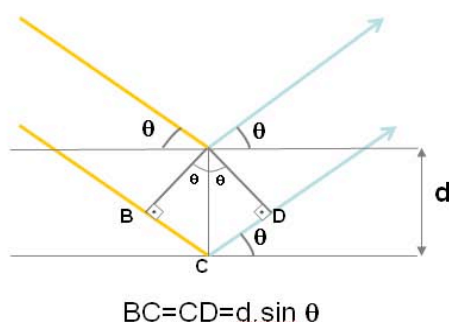


Figure 29 : Principe de la diffraction selon Bragg.
Les faisceaux de rayons X incidents sont représentés en jaune, ceux qui sont diffractés en bleu.

IV.1.2.3. Facteur de structure et densité électronique

Chaque tache hkl du cliché de diffraction représente une onde résultante de la somme des diffractions contributives de tous les électrons appartenant à la famille de plans réticulaires qu'elle représente. La fonction mathématique décrivant cette onde est appelée le **Facteur de structure F_{hkl}**. Il fait intervenir le pouvoir diffusant de chaque atome (f_j) de la molécule de la maille de coordonnées réduites x_j , y_j , z_j ainsi que sa position dans l'espace par rapport à un atome de base choisi comme origine :

$$F_{hkl} = \sum_j f_j e^{2i\pi(hx_j + ky_j + lz_j)} \quad (1)$$

Ce facteur de structure rassemble l'information sur l'amplitude $|F_{hkl}|$, qui est proportionnelle à la racine carrée de l'intensité mesurée, et la phase ϕ_{hkl} de l'onde diffusée suivant la relation:

$$F_{hkl} = |F_{hkl}| e^{i\phi_{hkl}}$$

En outre, le facteur de structure F_{hkl} correspond à la transformée de Fourier de la densité électronique $\rho(x,y,z)$. On peut ainsi avoir accès à la densité électronique en tout point du cristal par l'équation suivante, V étant le volume de la maille :

$$\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2i\pi(hx+ky+lz)}$$

Lors d'une expérience de diffraction, les clichés obtenus permettent de déterminer les intensités de chaque réflexion hkl. Chaque intensité mesurée est proportionnelle au carré du module du facteur de structure $|F_{hkl}|$ (correspondant à l'amplitude). Afin de déterminer le facteur de structure et ainsi remonter à la densité électronique, il nous manque donc l'information concernant la phase qui ne peut être connue lors d'une simple expérience de diffraction. Par conséquent, la densité électronique ne peut pas être calculée directement à partir des clichés de diffraction..

IV.1.3. Détermination des intensités de diffraction

Lors de l'enregistrement des clichés de diffraction d'un cristal, nous observons sur le détecteur des taches de diffraction : il s'agit des pixels du détecteur qui ont été excités par des rayons diffractés. Le nombre de photons reçus par l'ensemble des pixels formant la tache permet de déterminer l'intensité de cette dernière. Les clichés de diffraction obtenus sont donc traités afin d'extraire les valeurs d'intensité associées à chaque indice de Miller hkl . Ceci est effectué avec le groupe de programmes XDS (Kabsch, 1993).

La première étape du traitement des données de diffraction est l'**indexation**. Durant cette étape, la répartition des taches de diffraction est analysée ; des vecteurs entre les taches de diffraction sont mesurés ce qui permet de déterminer la maille du cristal.

L'étape d'**intégration** permet, quant à elle, de déterminer l'intensité des réflexions. Pour ce faire, un masque est défini autour de chaque tache de diffraction. La partie intérieure de celui-ci correspond à la partie à intégrer tandis que la partie extérieure correspond au bruit de fond qui sera retranché à l'intensité de la réflexion.

De par la présence d'opérations de symétrie dans le cristal, certaines réflexions sont reliées entre elles par symétrie et sont équivalentes. Elles devraient donc avoir la même intensité. L'étape de **mise à l'échelle** des données permet de moyenner les intensités de ces taches.

Au cours du traitement des clichés de diffraction, les programmes donnent différents facteurs statistiques permettant de contrôler la qualité de diffraction du cristal. La qualité du jeu de données peut être estimée par une valeur du **facteur R_{sym}** qui permet de comparer l'intensité de chaque réflexion équivalente par la symétrie du cristal à la valeur moyenne de ces réflexions :

$$R_{sym} = \frac{\sum_{hkl} |I_{hkl} - \langle I \rangle|}{\sum_{hkl} I_{hkl}}$$

Ainsi, le R_{sym} représente le désaccord entre les intensités des réflexions équivalentes par symétrie : plus il est bas, meilleur est l'accord. Si un jeu de données est indexé dans le mauvais groupe d'espace, les opérateurs de symétrie appliqués ne sont pas les bons, le facteur R_{sym} augmente très sensiblement car les réflexions considérées comme équivalentes par symétrie ne le sont pas. Au cours de l'enregistrement, lorsque le cristal se dégrade, l'intensité

diffractée diminue et la valeur du facteur R_{sym} augmente. On définit également un **rapport signal sur bruit** : $I/\sigma(I)$ qui permet de donner une estimation de l'intensité moyenne des réflexions mesurées en fonction de la résolution, où I est l'intensité des réflexions et $\sigma(I)$, l'écart type des intensités des réflexions (bruit de fond). Ainsi, plus le signal est élevé par rapport au bruit de fond, meilleures sont les estimations des intensités et donc des facteurs de structure.

A ce stade, seules les modules du facteur de structure des réflexions hkl sont connues. Pour remonter à la densité électronique, il nous faut donc déterminer les phases.

IV.1.4. Résolution du problème de phase

Pour calculer la carte de densité électronique de la macromolécule, deux éléments essentiels doivent être déterminés : les modules du facteur de structure des réflexions hkl et les phases. Cependant lors d'une expérience de cristallographie, seules les intensités (et donc les modules des facteurs de structure) des faisceaux diffractés sont enregistrés et l'information concernant chacune des phases des ondes diffractées est perdue. Les deux techniques utilisées pour déterminer ces phases et que je présenterai dans ce manuscrit sont la diffusion anormale et le remplacement moléculaire.

Dans le cas de la diffusion anormale, les propriétés particulières d'un atome lourd excité dans son seuil d'absorption sont utilisées pour déterminer sa position possible au sein de la structure cristalline, puis pour déterminer les phases globales du cristal.

Le remplacement moléculaire peut être utilisé dans le cas d'un cristal constitué d'une protéine dont la structure est proche d'une molécule de structure connue. On utilise cette structure comme modèle pour la détermination des phases.

IV.1.4.1. Préambule : la loi de Friedel

Dans un cristal natif, les intensités des réflexions (h, k, l) et $(-h, -k, -l)$ sont égales et leurs phases sont opposées. On appelle paires de Friedel, les facteurs de structure des deux réflexions h, k, l et $-h, -k, -l$. Ceci peut être représenté graphiquement sur le diagramme d'Argand de la figure 30. Si on considère les facteurs de structure de la protéine F_p , du diffuseur anormal F_a et de la structure totale F ainsi que les paires de Friedel de chaque facteur

de structure appelées F^+ , correspondant à $F(h, k, l)$, et F^- , correspondant à $F(-h, -k, -l)$, le cas d'une structure dérivée lors de la diffraction classique peut être représenté dans le diagramme suivant:

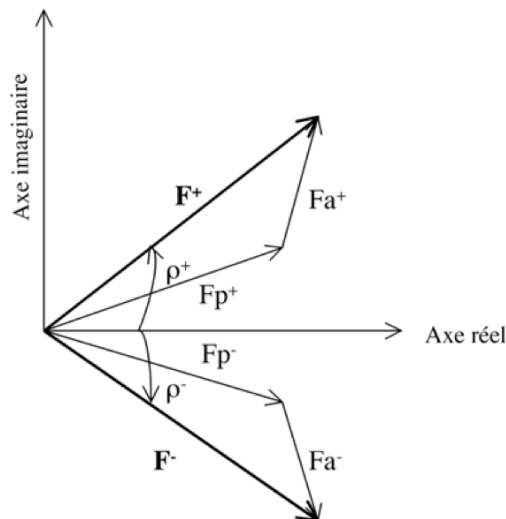


Figure 30 : Diagramme d'Argand montrant la loi de Friedel. Les facteurs de structure de la protéine F_p , du diffuseur anomal F_a et de la structure totale F sont représentés. ρ^+ et ρ^- correspondent aux phases de F^+ et F^- respectivement.

IV.1.4.2. La diffusion anormale

Généralités

Lorsque la longueur d'onde utilisée correspond au seuil d'absorption d'un atome lourd dans le cristal, il y a alors absorption du rayonnement incident par les électrons des couches profondes de l'atome lourd. Ces électrons se retrouvent excités. Le rayonnement diffracté par le diffuseur anomal (l'atome lourd) est déphasé par rapport aux autres rayonnements diffractés par les atomes « légers ». Pour tenir compte du déphasage, le facteur de structure de l'atome lourd ne vaut plus f^0 mais s'écrit :

$$f = f^0 + f'(\lambda) + if''(\lambda)$$

f^0 est indépendant de la longueur d'onde contrairement à f' et f'' . Lorsque le rayonnement incident est loin du seuil d'absorption de l'atome lourd, f' et f'' sont négligeables. En revanche, si la longueur d'onde est proche du seuil d'absorption d'un atome lourd, f' et f'' prennent des valeurs significatives. f' et f'' sont composées d'une partie réelle et d'une partie imaginaire.

En présence de diffusion anormale, la loi de Friedel, selon laquelle les modules des facteurs de structure $F(h, k, l)$ et $F(-h, -k, -l)$ sont égaux, est violée. En effet, dans le cas de la diffraction anormale, la structure complète du facteur de structure du diffuseur anormal possède trois composantes dont la partie imaginaire possède un déphasage de $\pi/2$. Les paires de Friedel $F(h, k, l)$ et $F(-h, -k, -l)$ ne sont donc plus équivalentes; leur intensité est en particulier différente.

Détermination des sous-structures des diffuseurs anomaux

En présence de diffuseurs anomaux, les deux réflexions symétriques ne sont plus équivalentes (du fait des contributions anormales f' et f''), et notamment, leurs intensités sont différentes. Cette différence d'intensité est plus généralement présente sur les paires de Bijvoet : ce sont les paires de réflexions symétriques dont les phases des facteurs de structure sont opposées. Les diffuseurs anomaux permettent de différencier ces réflexions symétriques, et apportent ainsi une information sur la phase du facteur de structure $F(hkl)$.

La différence entre les modules des facteurs de structure peut être calculée pour chaque paire de Bijvoet, appelée la différence de Bijvoet $\Delta F = |F^+| - |F^-|$. Cette différence est généralement relevée à la longueur d'onde correspondant au maximum de f'' , là où la différence est la plus grande. On peut alors calculer une carte de Patterson avec ΔF^2 comme coefficient. Cette carte correspond approximativement à une fonction d'auto-corrélation des diffuseurs anomaux qui présentent des maxima aux points correspondant aux distances interatomiques, permettant de déterminer la position des atomes lourds dans la structure.

Détermination graphique de la phase

Lorsque l'unité asymétrique contient un atome nettement plus lourd que les autres, la contribution de cet atome au facteur de structure est importante et souvent déterminante. Afin de rendre compte de la contribution de l'atome lourd, le facteur de structure total (F_T) peut s'écrire :

$$F_T(hkl) = F_P(hkl) + F_A(hkl)$$

avec F_P correspondant au facteur de structure de la protéine et F_A à celui de l'atome lourd.

Le facteur de structure de la protéine F_P recherché peut ainsi s'écrire :

$$F_P(hkl) = F_T(hkl) - F_A(hkl) \quad (2)$$

avec $F_A(hkl)$ préalablement déterminé en module et en phase.

Lors d'une expérience de diffusion anormale, un même cristal est utilisé sur lequel on enregistre plusieurs jeux de données à plusieurs longueurs d'onde différentes (méthode MAD). Trois ou quatre longueurs d'onde sont souvent choisies ; une correspondant au maximum de f'' pour laquelle la différence entre les F^+ et les F^- est la plus grande, une correspondant au longueur d'onde au minimum de f'' , et une ou deux longueurs d'ondes loin du seuil d'absorption du diffuseur qui permet de déterminer le module du facteur de structure de la protéine native. On peut considérer qu'aux deux premières longueurs d'onde, le diffuseur anormal se comporte comme deux dérivés différents.

L'équation (2) peut être représentée sur le diagramme d'Argand suivant :

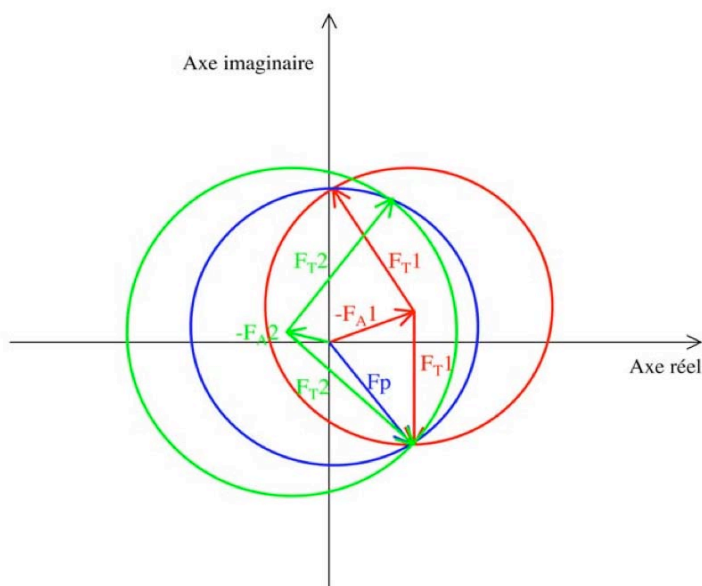


Figure 31 : Diagramme d'Argand montrant la détermination graphique de la phase par la méthode MAD. F_P correspond au facteur de structure de la protéine, F_A à celui du diffuseur anormal et F_T à celui de la structure totale.

Le module du facteur de structure F_P correspondant à la protéine est connu par la mesure des intensités de diffraction à la longueur d'onde loin du seuil d'absorption du diffuseur anormal. Un cercle de rayon correspondant au module de F_P (cercle bleu) peut donc être tracé. La fonction de Patterson permet de déterminer précisément la position des diffuseurs anomaux, nous permettant d'avoir accès au module et à la phase de F_A (en rouge). A une longueur d'onde de l'expérience, on connaît la valeur du module du facteur de structure de la protéine en complexe avec le diffuseur anormal F_{T1} (cercle rouge). L'intersection entre les

cercles bleus et rouges donnent deux solutions pour la phase de F_p . L'utilisation de la seconde longueur d'onde permet de déterminer la valeur du module de F_{T2} (cercle vert), et de lever l'ambiguïté en ce qui concerne la phase de F_p (à l'intersection des cercles bleu, rouge et vert).

IV.1.4.3. Le remplacement moléculaire

La méthode de phasage par remplacement moléculaire est employée lorsqu'on dispose d'une structure cristallographique d'une protéine possédant typiquement plus de 25 % d'identité de séquence avec la protéine d'intérêt. L'homologie de séquence n'est pas le seul critère pour résoudre une structure par remplacement moléculaire; c'est surtout l'homologie structurale qui permet d'utiliser les phases de la protéine modèle pour résoudre la structure de la protéine inconnue.

La méthode du remplacement moléculaire consiste à calculer les fonctions de Patterson, représentant l'ensemble des vecteurs interatomiques d'une molécule, pour le modèle cristallographique et pour la structure recherchée. La première étape consiste à déterminer les trois angles d'Euler, pour lesquels la fonction de corrélation entre la fonction de Patterson du modèle orienté et celle issue de la structure inconnue est maximale : c'est la fonction de rotation. Après avoir déterminé l'orientation du modèle, la deuxième étape consiste à repositionner le modèle par translation pour donner le meilleur accord entre les facteurs de structure calculés et mesurés. Les phases sont ainsi extraites. Le modèle obtenu de la protéine est néanmoins biaisé par le modèle de départ. Le modèle doit ensuite être affiné.

IV.1.5. Construction du modèle et affinement

À partir des phases expérimentales, déterminées soit par la méthode MAD, soit par remplacement moléculaire, on dispose d'une carte de densité électronique dans laquelle on peut construire un modèle de la structure recherchée. Des programmes, comme ARP/WARP (Morris *et al.*, 2003), permettent une construction automatique si les phases et la résolution sont de bonne qualité (résolution d'au moins 2,5 Å). La procédure d'affinement permet de corriger le modèle initial afin d'obtenir un accord optimal entre les modules des facteurs de structure observés et ceux calculés à partir du modèle construit. En pratique, on procède en alternant des étapes de reconstitution manuelle du modèle avec des étapes d'affinement

numérique réalisé avec le logiciel REFMAC5 (Murshudov *et al.*, 1997). Ce dernier affine les paramètres x, y, z et le facteur d'agitation thermique de chaque atome, tout en imposant certaines contraintes géométriques et stéréochimiques, comme les distances et les angles de liaison.

Afin de suivre l'affinement du modèle, deux grandeurs caractéristiques de l'état d'avancement de l'affinement sont utilisées, les **facteurs R et R_{free}** que l'on essaie de minimiser.

$$R = \frac{\sum_{hkl} \|F_{obs} - k|F_{calc}|\|}{\sum_{hkl} |F_{obs}|}$$

Le facteur R, faisant partie de la fonction à minimiser, n'est pas complètement objectif et diminue nécessairement au cours de l'affinement puisque l'on optimise l'accord entre facteur de structure observé et calculé. C'est donc le facteur R_{free}, moins biaisé, qui est utilisé puisqu'il est calculé sur un ensemble de réflexions exclues de la procédure d'affinement. Ces réflexions sont choisies de manière aléatoire et représentent le plus souvent 5 % des réflexions totales. Une diminution du facteur R sans diminution corrélée du facteur R_{free} indique un affinement inefficace qui consiste à faire entrer artificiellement dans le modèle le bruit des cartes de densité électronique.

IV.1.6. Analyse des structures

Afin d'évaluer et de valider un modèle, plusieurs critères doivent être surveillés :

- Lors du traitement des données : la résolution, R_{sym} (ou R_{merge}), le rapport signal sur bruit I/σ(I) et la complétude.
- Lors de la construction et de l'affinement du modèle : les statistiques d'affinement R et R_{free}, les déviations moyennes des distances et les angles entre les liaisons.
- pour le modèle final, une structure est estimée correcte quand les écarts moyens n'excèdent pas 2° pour les angles et 0,02 Å pour les distances. La stéréochimie de la chaîne polypeptidique est estimée par le diagramme de Ramachandran.

IV.2. Cristallogenèse et cristallographie de SrtC-1

IV.2.1. Cristallogenèse de SrtC-1

La méthode utilisée est celle de la diffusion de vapeur en goutte suspendue réalisée à 20°C. En utilisant la plateforme de cristallisation à haut débit disponible à l'EMBL, des cristaux ont été obtenus dans une condition du kit Crystal Screen I de Hampton (0,1 M Tris pH 8,5, 0,2 M MgCl₂ hexahydrate, 30% PEG4000). La condition a alors été reproduite à la main en mixant 1 µl de protéine purifiée (à environ 10 mg/ml) avec 1 µl de solution A contenant 0,1 M Tris pH 8,5, 0,2 M MgCl₂ hexahydrate, 30 % PEG4000. Les cristaux obtenus sont ensuite plongés successivement dans une solution cryoprotectrice contenant la solution A supplémentée de concentration croissante de glycérol (2%, 5%, 10%, 15%) avant d'être refroidis dans l'azote liquide. Les cristaux de SrtC-1 sélénée ont également été obtenus dans la même condition.

IV.2.2. Enregistrement des données de diffraction

Un premier jeu de données natif a été enregistré sur la ligne de lumière ID14-EH3 de l'ESRF. Un jeu de données MAD fut ensuite enregistré sur la ligne de lumière ID23-EH1 de l'ESRF, permettant de modifier les longueurs d'onde pendant l'expérience. Un spectre de fluorescence, au seuil d'absorption du sélénium, a été enregistré sur un cristal de SrtC-1 sélénée. L'analyse de ce spectre permet de déterminer les variations des coefficients f et f'' en fonction des variations de l'énergie des photons incidents. La longueur d'onde correspondant au seuil d'absorption du sélénium pour lequel la valeur de f'' est maximale a été choisie comme longueur d'onde du pic à 0,98055 Å. La seconde longueur d'onde utilisée correspond au minimum de f'' (inflection) à 0,980750 Å. Une troisième ainsi qu'une quatrième longueur d'onde de références ont été choisies à 0,982450 Å et 0,978550 Å.

IV.2.3. Structure de SrtC-1

Les différents jeux de données ont été traités avec le groupe de programmes XDS (Kabsch, 1993). Les données correspondant aux 4 longueurs d'ondes enregistrées sur SrtC-1

sélénée ont été mises ensemble en utilisant le programme CAD de CCP4 (Collaborative Computational Project, 1994). Le phasage a été réalisé par la méthode de dispersion anormale MAD en utilisant le signal du sélénium. 14 sites de fixation de sélénium ont été localisés dans l'unité asymétrique par le programme AutoSHARP (Bricogne *et al.*, 2003). Après une étape d'aplatissement de solvant, ARP/wARP (Perrakis *et al.*, 1999) a construit automatiquement dans la carte de densité électronique 338 résidus sur les 426 résidus attendus.

Afin d'améliorer la précision du modèle et étant donné que nous disposons d'un jeu de données natif à une meilleure résolution (1,24 Å), nous avons effectué un remplacement moléculaire. Pour cela, nous avons injecté les phases du modèle construit précédemment dans le jeu de données collectées à 1,24 Å. Le modèle obtenu a été ensuite affiné par alternance de modélisation manuelle avec le programme COOT (Emsley and Cowtan, 2004) et de cycles d'affinement numérique grâce au programme REFMAC5 (Murshudov *et al.*, 1997). L'affinement s'est poursuivi jusqu'à obtenir des paramètres stéréochimiques corrects en terme d'angles et de longueurs de liaisons atomiques, ainsi que des données statistiques satisfaisantes en terme de facteur R.

IV.3. Cristallogenèse et cristallographie de SrtC-3

IV.3.1. Cristallogenèse de SrtC-3

L'approche utilisée pour obtenir des cristaux de SrtC-3 est celle de la méthode de criblage manuel grâce aux différents kits Hampton disponibles au laboratoire. La méthode de diffusion de vapeur en goutte suspendue a été effectuée. Après avoir tester 144 conditions, une seule a permis de mettre en évidence une piste prometteuse pour l'obtention de cristaux. Cette condition a été légèrement modifiée et un cristal a été obtenu dans 0,1 M Mes pH 5, 0,5 M (NH₄)₂SO₄, 1 M LiSO₄ à 20°C. Étant donné la haute concentration en lithium présente dans la solution de cristallogenèse, nous n'avons effectué aucune cryoprotection préalable du cristal. Le cristal a été refroidi directement dans l'azote liquide.

IV.3.2. Structure de SrtC-3

Après avoir enregistré un jeu de données au Laboratoire de Cristallogenèse et de Cristallographie des Protéines à l'IBS sur un générateur à anode tournante, ces données ont

été indexées et intégrées à l'aide du groupe de programmes XDS (Kabsch, 1993). La structure de SrtC-3 a été résolue par la méthode de remplacement moléculaire en utilisant le programme PHASER (McCoy *et al.*, 2005; Storoni *et al.*, 2004). Le modèle utilisé pour le remplacement moléculaire est celui qui correspond au cœur de la structure de SrtC-1, c'est-à-dire des résidus 82 à 209, dans lesquels les acides aminés qui diffèrent entre SrtC-1 et SrtC-3 ont été remplacés par des alanines. La construction automatique du modèle ainsi que l'affinement ont été réalisés comme décrit précédemment avec SrtC-1.

IV.4. Cristallogénèse et cristallographie de RrgB

IV.4.1. Cristallogénèse de RrgB

Les essais de cristallogénèse de RrgB ont débuté par un criblage manuel de 192 conditions avec les différents kits Hampton disponibles au laboratoire, en utilisant la méthode de diffusion de vapeur en goutte suspendue. 1 μ l de protéine purifiée à environ 10 mg/ml et 1 μ l de solution de cristallisation sont mélangées. Après plus d'un mois, une piste prometteuse pour l'obtention de cristaux a été décelée et, après affinement des conditions de cristallisation, des cristaux sont apparus dans: 0,1 M Tris pH 7,5, 18% PEG12000. Ces cristaux poussent en un peu moins de deux mois. Des cristaux de RrgB sélénée ont été obtenus dans 0,1 M Tris pH 7,5, 23% PEG8000.

IV.4.2. Enregistrement des données de diffraction

Un jeu de données à 3,7 Å a été enregistré sur la ligne de lumière ID14-EH1 de l'ESRF, avec un pas d'oscillation de 0,5°. Ce jeu a été indexé et intégré par le groupe de programmes XDS (Kabsch, 1993).

Des données de diffraction sur un cristal de protéine sélénée ont été enregistrées à la longueur d'onde correspond au pic du sélénium (0,97950 Å) à 3,1 Å. Un pas d'oscillation de 0,1° a été effectué pour éviter le recouvrement des taches de diffraction. Ce jeu a été indexé et intégré par le groupe de programmes XDS (Kabsch, 1993).

Résultats et Discussion

I. Etude fonctionnelle des composants du pilus de *S. pneumoniae*

L'ensemble des résultats de cette partie a fait l'objet d'un article (Manzano *et al.*, 2008).

I.1. Rôle des sortases SrtC-1, SrtC-2 et SrtC-3 dans la polymérisation de RrgB

I.1.1. Analyse *in vitro*

Chez *S. pneumoniae*, le corps du pilus est composé de monomères de RrgB qui sont liés covalamment entre eux. Des études précédentes, réalisées *in vivo*, ont montré qu'il était possible de visualiser les pili, issus de la surface du pneumocoque, sur un gel dénaturant SDS par la présence de plusieurs bandes de haut poids moléculaires. Ceci est en effet rendu possible puisque les sous-unités de RrgB sont liées les unes aux autres par des liaisons covalentes et ne sont donc pas brisées par l'action du chauffage et du SDS (Barocchi *et al.*, 2006 ; Hilleringmann *et al.*, 2008). Nous sommes alors partis de ce principe pour initier notre étude sur le processus de formation du corps du pilus et tester la capacité de chaque sortase à polymériser RrgB. Après expression et purification simultanées de SrtC-1, SrtC-2, SrtC-3 et RrgB, chaque sortase est incubée avec RrgB à 37°C pendant une nuit. La capacité de former des fibres de RrgB *in vitro* est ensuite analysée en recherchant des bandes de plus haut poids moléculaire sur un gel de gradient SDS. Pour plus de précision, un Western blot est réalisé avec les anticorps dirigés contre RrgB.

Plusieurs bandes de haut poids moléculaire sont détectées quand RrgB est incubée avec SrtC-1 (figure 32, piste 2). En revanche, lorsque RrgB est incubée en présence de SrtC-2, aucune polymérisation de RrgB n'est identifiée (figure 32, piste 3). Il est cependant à noter que deux faibles bandes reconnues par les anticorps anti-RrgB sont visibles avec SrtC-3 (figure 32, piste 4, indiquées par des étoiles). Ces résultats suggèrent donc que, *in vitro*, SrtC-1 est la sortase la plus efficace pour polymériser RrgB, bien que SrtC-3 soit capable de polymériser RrgB, mais de façon modérée.

Afin de vérifier que la polymérisation de RrgB détectée sur gel SDS est bien le fruit d'une activité sortasique, nous avons testé la capacité du mutant SrtC-1C193A, dans lequel la cystéine catalytique a été remplacée par une alanine, à former les fibres de RrgB. Les sortases étant des enzymes qui utilisent leur cystéine pour catalyser la réaction de transpeptidation, une perte de ce résidu entraînerait donc une perte de leur activité, comme ce fut montré pour SrtA de *S. aureus* (Ton-That *et al.*, 2000, 2002). Après incubation avec SrtC-1C193A, RrgB reste monomérique et aucune bande de haut poids moléculaire n'est présente (figure 32, piste 5). En conclusion, la polymérisation de RrgB *in vitro* est bien le résultat de l'activité de SrtC-1.

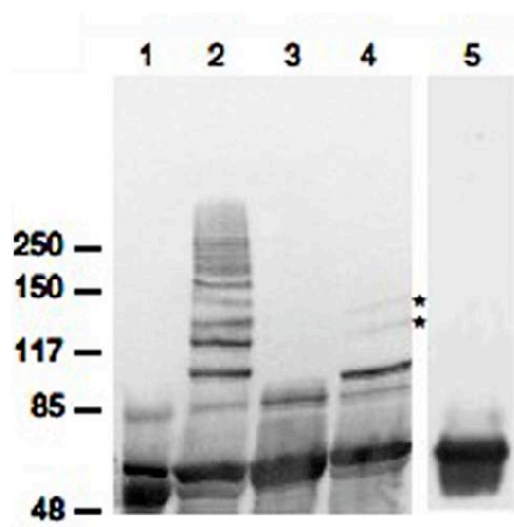


Figure 32 : Analyse *in vitro* de la formation de fibres de RrgB catalysée par les sortases pneumococciques (Manzano *et al.*, 2008). Western Blot révélé avec des anticorps anti-RrgB. (1) RrgB monomérique. (2) RrgB + SrtC-1. (3) RrgB + SrtC-2. (4) RrgB + SrtC-3. (5) RrgB + SrtC-1C193A. Les étoiles indiquent la présence de deux faibles bandes protéiques.

1.1.2. Analyse *in vivo*

Afin de confirmer les résultats obtenus *in vitro*, nous avons étudié, *in vivo*, l'effet de la perte d'activité de SrtC-1. Pour cela, dans la souche pneumococcique TIGR4, le gène codant pour SrtC-1 (gène *sp0466* de l'îlot *rlrA*) a été délété et le phénotype analysé. Cette étude a été réalisée par Anne-Marie Di-Guilmi et Lamya El Mortaji à l'IBS. Afin de récupérer les pili ancrés au peptidoglycane, les bactéries mutantes TIGR4 Δ *sp0466* et celles sauvages ont été traitées au lysozyme et à la mutanolysine, agents qui hydrolysent le peptidoglycane. Une centrifugation subséquente a permis de se débarrasser des agrégats cellulaires et de ne garder

que les composants ancrés au peptidoglycane, dont les pili. Les extraits totaux (c'est-à-dire avant centrifugation) ainsi que les surnageants ont été alors chargés sur gel SDS et la présence de pili a été analysée par western blot avec les anticorps anti-RrgB.

La souche sauvage TIGR4 possède des pili à sa surface comme en témoigne la présence de bandes de haut poids moléculaire sur gel SDS (figure 33, pistes 1 et 2). La souche mutante, n'exprimant plus la sortase SrtC-1, possède quant à elle une quantité plus faible de pili à sa surface, ce qui est en accord avec l'augmentation de la quantité de RrgB monomérique (figure 33, pistes 3 et 4). SrtC-1 apparaît donc bien être la sortase principale polymérisant le pilus, bien que son absence puisse être partiellement compensée par d'autre(s) sortase(s) comme en témoigne la présence d'une faible quantité de pili détectée à la surface du mutant TIGR4 Δ sp0466.

Il est à noter que les pili, obtenus *in vivo* et observés sur gel, possèdent une taille supérieure à celle des fibres que nous sommes capables de former *in vitro*. Cela peut s'expliquer par le fait que les pili extraits du pneumocoque sont additionnés de fragments de peptidoglycane, rendant ainsi leur masse plus grande. De plus, il se peut également que des protéines se retrouvent ancrées sur le peptidoglycane. La masse des pili observée *in vivo* reflète donc, en réalité, la masse des pili additionnée de celle des fragments de peptidoglycane sur lesquels peuvent se localiser des protéines.

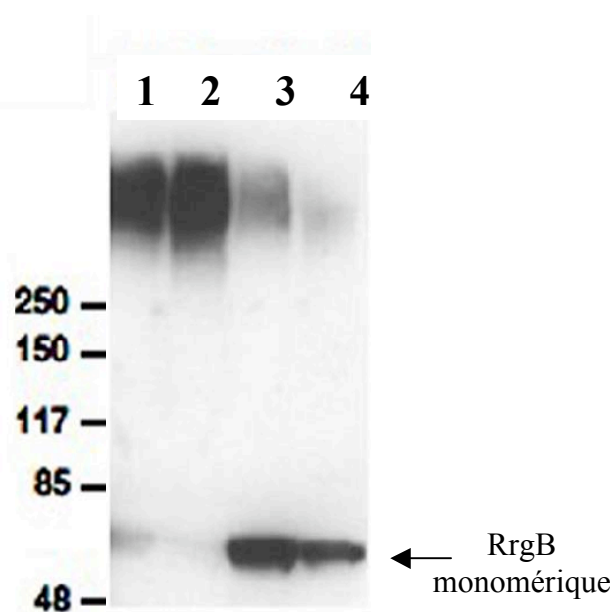


Figure 33 : Analyse de la délétion du gène *sp0466* codant pour SrtC-1 dans la souche pneumococcique TIGR4 (Manzano *et al.*, 2008). Western Blot développé avec des anticorps dirigés contre RrgB. Après traitement à la mutanolysine et au lysozyme, les échantillons avant et après centrifugation sont déposés. (1) extrait total de TIGR4. (2) surnageant de TIGR4. (3) extrait total de TIGR4 Δ sp0466. (4) surnageant de TIGR4 Δ sp0466.

En conclusion, ces résultats *in vivo* confirment ce qui a été obtenu *in vitro* à savoir que SrtC-1 joue un rôle clé dans le processus de polymérisation du pilus. L'absence de cette enzyme peut être toutefois compensée par une activité de polymérisation de RrgB réalisée par l'une ou l'autre des sortases (ou encore les deux), ce qui a été également établi par les travaux de LeMieux et ses collègues. En effet, ceux-ci ont montré que des souches pneumococques, dans lesquelles une des sortases est manquante, est toujours capable d'exprimer des pili à sa surface mais en quantité inférieure (LeMieux *et al.*, 2008).

I.2. Caractérisation des fibres de RrgB formées *in vitro*

I.2.1. Purification des fibres

Après incubation à 37°C pendant une nuit, le mélange SrtC-1/RrgB, réalisé à grande échelle, est chargé sur une colonne d'exclusion de taille afin de séparer les fibres des monomères de SrtC-1 et RrgB. Le profil d'élution obtenu est montré ci-dessous.

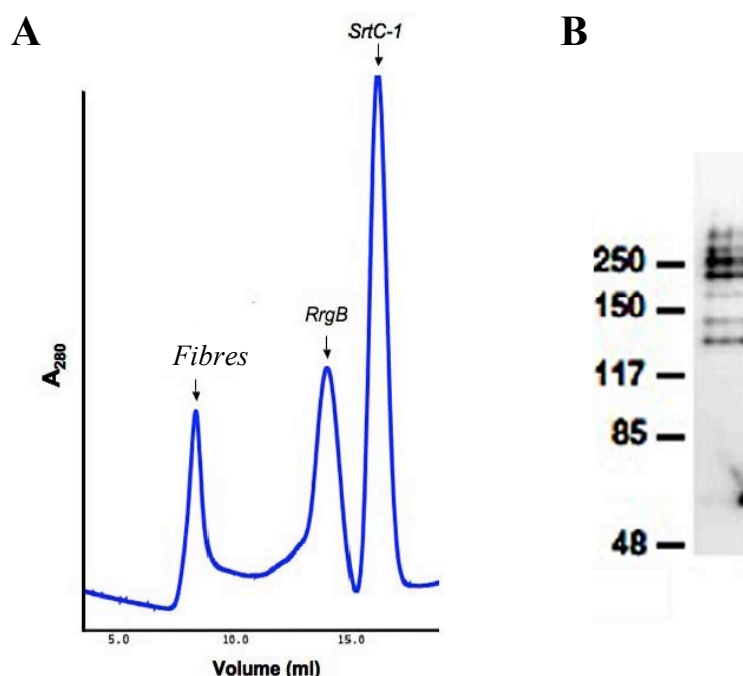


Figure 34 : Caractérisation des fibres de RrgB synthétisées *in vitro* (Manzano *et al.*, 2008). (A) Profil d'élution de la chromatographie d'exclusion du mélange SrtC-1/RrgB après incubation d'une nuit à 37°C. (B) Analyse par Western Blot, avec les anticorps anti-RrgB, de la fraction de la gel filtration correspondant au fibres. Cet échantillon sera ensuite analysé par microscopie électronique et par spectrométrie de masse.

Les fractions correspondant aux différents pics ont été analysées sur gel SDS. La fraction qui contient les fibres, montré en figure 34B, a été concentrée puis analysée par microscopie électronique. D'autre part, afin d'établir l'authenticité des bandes de haut poids moléculaire visualisées sur gel, celles-ci ont été analysées par spectrométrie de masse.

1.2.2. Spectrométrie de masse des fibres

Une partie de l'échantillon provenant de la gel filtration et contenant les fibres purifiées a été chargé sur gel SDS. Après coloration des bandes protéiques, celles-ci ont été découpées et analysées séparément par chromatographie liquide nano-débit couplée à un spectromètre de masse en tandem (nanoLC-MS/MS). Ces études ont été réalisées par la plateforme "EDyP-Service" disponible au CEA de Grenoble par l'équipe de Jérôme Garin.

Après digestion trypsique de chacune des bandes, les peptides issus de l'hydrolyse enzymatique des protéines vont dans un premier temps être séparés par chromatographie suivant leur hydrophobicité. Les peptides sont ensuite injectés séquentiellement dans la source du spectromètre de masse où ils sont ionisés et fragmentés. L'analyse des spectres MS/MS expérimentaux, résultants de la fragmentation des peptides, permettent ainsi l'identification des peptides issues de la digestion trypsique.

Les peptides issus de la digestion trypsique des différentes bandes protéiques appartiennent à RrgB, confirmant ainsi l'identité des bandes visualisées sur gel SDS (figure 35)

Masse observée	Masse expt	Masse calc	Δ	AA ₁	AA _x	Séquence	Protéine identifiée
513,7	1025,43	1025,48	-0,05	243	251	MTEGLAFNK	RrgB
523,74	1045,47	1045,56	-0,08	129	138	FNTANLPAAK	RrgB
563,24	1124,46	1124,57	-0,11	300	309	ITYSATLNDK	RrgB
588,79	1175,57	1175,62	-0,05	580	590	QPAGYALLTSR	RrgB
602,81	1203,61	1203,67	-0,06	377	387	VVQTVTLTDDK	RrgB
714,33	1426,65	1426,71	-0,06	243	255	MTEGLAFNKGTVK	RrgB
958,42	1914,84	1914,92	-0,08	410	427	GYSADYQEITTAGEIAVK	RrgB

Masse observée	Masse expt	Masse calc	Δ	AA ₁	AA _x	Séquence	Protéine identifiée
523,84	1045,67	1045,56	0,11	129	138	FNTANLPAAK	RrgB
563,35	1124,69	1124,57	0,12	300	309	ITYSATLNDK	RrgB
832,01	1662,01	1661,82	0,19	228	242	IPALANYATANWSDR	RrgB
566,01	1695,02	1694,85	0,17	388	402	NTVTVNGLDKNTEYK	RrgB
639,39	1915,14	1914,92	0,22	410	427	GYSADYQEITTAGEIAVK	RrgB

688,76	2063,25	2063,04	0,21	461	480	LAGAEFVIANADNAGQYLAR	RrgB
714,77	2141,30	2141,06	0,23	209	227	DTPVNHQVGDVVEYEIVTK	RrgB
749,49	2245,45	2245,22	0,24	377	397	VVQTVTLTDDKNTVTVNGLDK	RrgB
781,14	2340,39	2340,15	0,24	141	162	IYEIHSLSYVGEDGATLTGSK	RrgB
828,85	2483,54	2483,25	0,28	206	227	VDKDTPVNHQVGDVVEYEIVTK	RrgB

Masse observée	Masse expt	Masse calc	Δ	AA ₁	AA _x	Séquence	Protéine identifiée
588,87	1175,72	1175,63	0,09	580	590	QPAGYALLTSR	RrgB
832,01	1662,01	1661,82	0,20	228	242	IPALANYATANWSDR	RrgB
958,58	1915,14	1914,92	0,22	410	427	GYSADYQEITTAGEIAVK	RrgB
688,77	2063,29	2063,04	0,24	461	480	LAGAEFVIANADNAGQYLAR	RrgB
781,14	2340,41	2340,15	0,26	141	162	IYEIHSLSYVGEDGATLTGSK	RrgB
621,88	2483,51	2483,25	0,25	206	227	VDKDTPVNHQVGDVVEYEIVTK	RrgB

Masse observée	Masse expt	Masse calc	Δ	AA ₁	AA _x	Séquence	Protéine identifiée
958,58	1915,14	1914,92	0,22	410	427	GYSADYQEITTAGEIAVK	RrgB
688,76	2063,26	2063,04	0,22	461	480	LAGAEFVIANADNAGQYLAR	RrgB
781,13	2340,37	2340,15	0,22	141	162	IYEIHSLSYVGEDGATLTGSK	RrgB
857,17	2568,49	2568,20	0,29	42	65	LLATDGDMDKIANELETGNYAGNK	RrgB

Masse observée	Masse expt	Masse calc	Δ	AA ₁	AA _x	Séquence	Protéine identifiée
688,75	2063,24	2063,04	0,20	461	480	LAGAEFVIANADNAGQYLAR	RrgB

Masse observée	Masse expt	Masse calc	Δ	AA ₁	AA _x	Séquence	Protéine identifiée
588,88	1175,75	1175,63	0,12	580	590	QPAGYALLTSR	RrgB
958,57	1915,13	1914,92	0,21	410	427	GYSADYQEITTAGEIAVK	RrgB
688,76	2063,25	2063,04	0,21	461	480	LAGAEFVIANADNAGQYLAR	RrgB
781,14	2340,38	2340,15	0,24	141	162	IYEIHSLSYVGEDGATLTGSK	RrgB
861,18	2580,51	2580,23	0,28	591	615	QKFEVTATSYSATGQGIEYTAGSGK	RrgB

Masse observée	Masse expt	Masse calc	Δ	AA ₁	AA _x	Séquence	Protéine identifiée
523,82	1045,62	1045,56	0,06	129	138	FNTANLPAAK	RrgB
563,33	1124,65	1124,57	0,08	300	309	ITYSATLNDK	RrgB
588,88	1175,74	1175,63	0,11	580	590	QPAGYALLTSR	RrgB
714,43	1426,85	1426,71	0,14	243	255	MTEGLAFNKGTVK	RrgB

Figure 35 : Analyse des spectres expérimentaux MS/MS obtenus après digestion trypsique des bandes protéiques visualisées sur gel SDS à haut poids moléculaire (Manzano *et al.*, 2008).

Masse observée : masse du peptide sélectionné (ratio m/z).

Masse expt : Masse moléculaire expérimentale (Da) du peptide sélectionné.

Masse calc : Masse moléculaire calculée (Da) du peptide sélectionné.

Δ : Différence de masse (Da) entre Masse expt et Masse calc.

AA1: Position du premier acide aminé du peptide sur la protéine identifiée.

AAx: Position du dernier acide aminé du peptide sur la protéine identifiée.

Séquence: Séquence du peptide identifié.

I.2.3. Microscopie électronique des fibres

Après purification des fibres sur gel filtration, celles-ci ont été également étudiées en microscopie électronique à coloration négative par le groupe de Guy Schoehn à l'IBS. Deux populations de fibres, avec deux diamètres différents, sont présentes (figure 36). La première, majoritaire, présente un diamètre de 3,5 nm environ. Le second type de fibres a un diamètre plus grand, d'approximativement 7 nm. Ces dernières semblent être potentiellement formées par l'association latérale de deux fibres plus fines (figure 36, d et g). De plus, les fibres observées sont « coudées », suggérant ainsi une certaine flexibilité de ces structures (figure 36, i et j).

Le poids moléculaire des fibres observées peut être approximativement calculé, comme décrit dans la section Matériel et Méthodes. Ainsi, le poids moléculaire de la plus grande fibre a été estimé à 480 kDa et celui de la plus petite à 140 kDa environ.

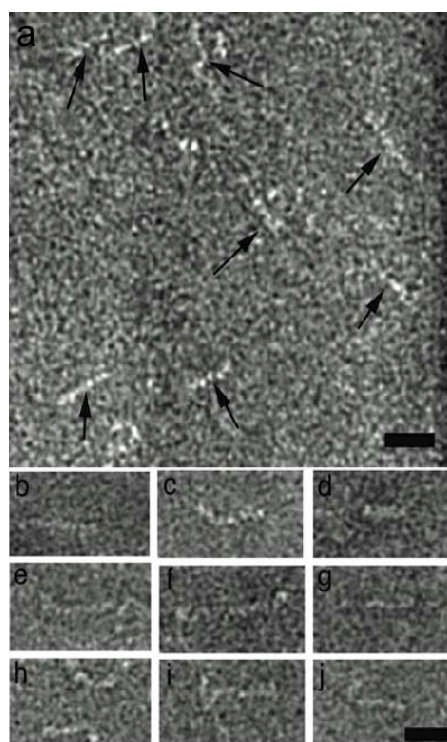


Figure 36 : Analyse de microscopie électronique en coloration négative des fibres de RrgB formées *in vitro* (Manzano *et al.*, 2008). Les flèches montrent les fibres dont le diamètre est d'environ 3,5 ou 7 nm. L'échelle représentée est de 20 X 5 nm. b-j : galerie de fibres.

In vivo, les pili de *S. pneumoniae* ont été décrits comme des structures flexibles formées par de fins protofilaments arrangés en une superstructure « coiled-coil » (Hilleringmann *et al.*, 2008) (figure 37A). Ainsi, deux protofilaments de 3,5 nm de diamètre se croisent étroitement pour former des pili de diamètre 6,8 nm. Des zones, dans lesquelles

l'intersection entre les deux protofilaments est plus relâchée, se retrouvent également sur le pilus produisant alors un diamètre de 9,5 nm (Hilleringmann *et al.*, 2008) (figure 37B).

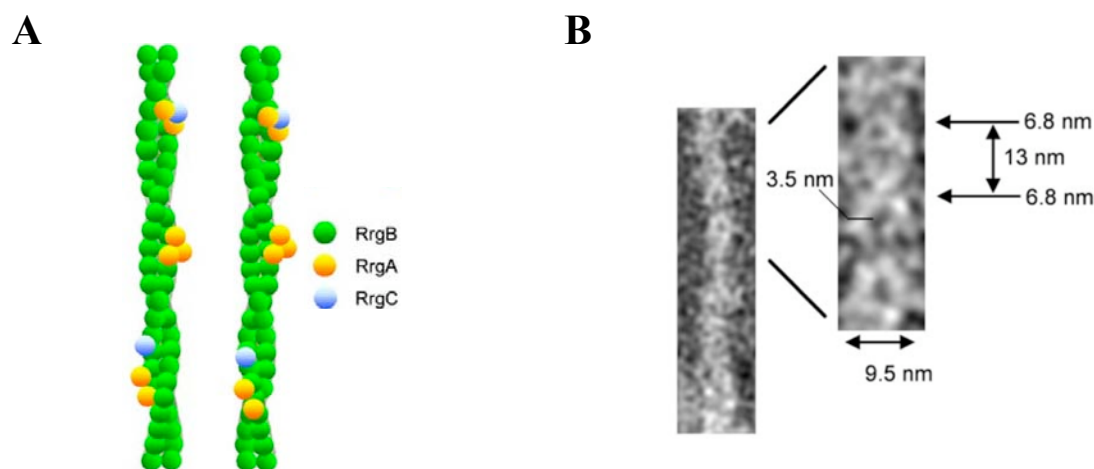


Figure 37 : Structure des pili de *S. pneumoniae*. A/ Modèle du pilus de *S. pneumoniae*. Deux protofilaments, composés de RrgB polymérisée, s'assemblent en une structure coiled-coil sur laquelle sont localisés les protéines RrgA et RrgC (Hilleringmann *et al.*, 2008). B/ Images de cryomicroscopie électronique d'un pilus pneumococcique. Le diamètre du protofilament est de 3,5 nm, celui du pilus de 6,8 nm (et peut atteindre 9.5 nm de diamètre) (Hilleringmann *et al.*, 2008).

Par conséquent, les fibres que nous sommes capables de produire *in vitro* ont approximativement le même diamètre que les protofilaments de 3,5 nm et les filaments de 6,8 nm identifiés à la surface du pneumocoque.

L'assemblage formé *in vitro*, d'une longueur moyenne de 35 nm, est cependant plus petit que les pili visualisés *in vivo* qui ont une longueur de 500 μm à plus de 1 μm , comme le montre la figure 38. Il a été montré que les pili présents à la surface de diverses bactéries à Gram-positif possèdent des longueurs variables. Ainsi, des formes assez courtes de pili ont pu être identifiées chez *S. salivarius*, *S. agalactiae*, *E. faecalis* ou encore *S. pneumoniae*. Il est donc tout à fait possible que les structures que nous formons *in vitro* représentent ces « formes » minimales.

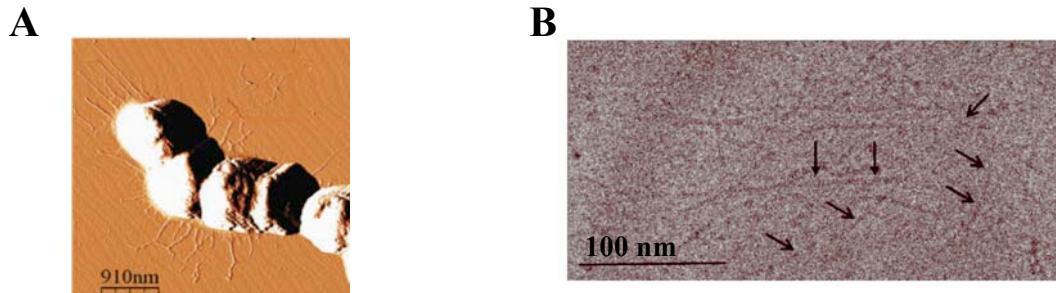


Figure 38 : Le pneumocoque et ses pili. A/ Souche pneumococcique TIGR4 possédant des pili à sa surface observée par microscopie à force atomique (Falker *et al.*, 2008). B/ Image de cryo-microscopie électronique de pili purifiés de la surface du pneumocoque. Les flèches ouvertes et fermées indiquent la présence de différentes tailles de pili (Hilleringmann *et al.*, 2008).

Une autre hypothèse peut cependant être envisagée. Il est effectivement possible que les fibres formées *in vitro* soient courtes car SrtC-1 ne polymérise qu'inefficacement RrgB (ce qui est démontré par le nombre fini de bandes présentes sur gel SDS) et que cette enzyme a besoin d'autres composants du pilus pour former une fibre entière, telles que les autres sortases. Cette possibilité a été testée *in vitro* en incubant RrgB avec les trois sortases SrtC-1, SrtC-2 et SrtC-3 mais aucune différence n'a été observée. Par conséquent, la formation *in vivo* des pili doit probablement être un processus très bien orchestré qui requiert l'ensemble des composants impliqués dans la formation du pilus et associés à la membrane (aussi bien les sortases que les protéines Rrg) afin d'optimiser la reconnaissance des partenaires mais aussi la catalyse. Il est ainsi concevable qu'une telle organisation ne peut être reproduite *in vitro*, ce qui expliquerait la présence de courtes fibres dans notre cas.

Il est intéressant à noter que, pour la première fois, le rôle des sortases dans le processus de polymérisation du pilus a pu être mis en évidence *in vitro* en utilisant les protéines substrats.

II. Détermination de la structure des sortases SrtC-1 et SrtC-3

Le traitement des jeux de données ainsi que la résolution des deux structures des sortases SrtC-1 et SrtC-3 ont été réalisés avec l'aide de Carlos Contreras-Martel. Les deux structures ont été déposées dans la PDB sous le code de 2W1J pour SrtC-1 et 2W1K pour SrtC-3 et ont été publiées dans Manzano *et al.*, 2008.

II.1. Structure de SrtC-1

II.1.1. Production de SrtC-1

Seule la région du génome de *S. pneumoniae* TIGR4 codant pour les acides aminés 17 à 228 a été clonée (clonage réalisé par Anne-Marie Di-Guilmi). En effet, la partie N-terminale, assez hydrophobe, et la partie C-terminale, prédite comme un ancrage transmembranaire par le logiciel SMART (<http://smart.embl-heidelberg.de>), ont été déléetées afin de s'affranchir d'éventuels problèmes de solubilité.

La protéine SrtC-1 est exprimée de façon recombinante chez *E. coli* et purifiée relativement facilement avec un rendement de bonne qualité (20 à 30 mg/L de culture).

II.1.2. Cristallogénèse de SrtC-1

Des premiers cristaux ont été obtenus avec le robot Cartésien de l'EMBL dans 0,1 M Tris pH 8,5, 0,2 M MgCl₂ hexahydrate, 30% PEG4000. Ces cristaux, étaient trop petits pour poursuivre des études de cristallographie aux rayons X. Ils ont été reproduits manuellement et apparaissent entre 48 et 72 heures dans le tampon cité précédemment (figure 39).



Figure 39 : Cristaux de SrtC-1 dans 0.1 M Tris pH 8.5, 0.2 M MgCl₂ hexahydrate, 30% PEG4000.

Ces cristaux sont ensuite plongés successivement dans des solutions cryoprotectrices contenant une concentration croissante de glycérol (jusqu'à 15%) puis congelés par immersion dans l'azote liquide.

II.1.3. Production de cristaux de SrtC-1 sélénée

Le remplacement moléculaire, en utilisant la structure de SrtA de *S. aureus* comme modèle, n'a pas permis de résoudre la structure de SrtC-1. Nous avons alors décidé d'exprimer la protéine sélénée, dans laquelle les résidus méthionines sont remplacés par des sélénométhionines, et d'utiliser la méthode MAD comme méthode de phasage. Un protocole similaire à celui de SrtC-1 native a été utilisé pour purifier SrtC-1 sélénée. Cependant, 10 mM de DTT a été ajoutés dans le tampon de gel filtration. Les cristaux de SrtC-1 sélénée sont obtenus également dans 0,1 M Tris pH 8,5, 0,2 M MgCl₂ hexahydrate, 30% PEG4000. La cryoprotection est la même que pour les cristaux natifs de SrtC-1.

II.1.4. Résolution de la structure de SrtC-1

II.1.4.1. Collecte des données natives et sélénées

La collecte des données de diffraction sur un cristal de SrtC-1 native a été réalisée sur la ligne de lumière ID14-EH3 de l'ESRF, à une longueur d'onde de 0,93100 Å. La collecte des données de diffraction sur un cristal de SrtC-1 sélénée a été effectuée sur la ligne ID23-EH1 de l'ESRF. Les longueurs d'onde ont été choisies à l'aide des spectres de fluorescence expérimentaux (figure 40). Nous avons ainsi réalisé 4 collectes aux longueurs d'onde du pic (maximum de f'') : 0,98055 Å, de l'inflexion (minimum de f'') : 0,980750 Å, et des références : 0,982450 Å et 0,978550 Å.

Pour chacun des jeux de données, 180 images, par incrémentation de 1°C, ont été enregistrées, permettant ainsi une bonne redondance des pics de diffraction et une bonne complétude (tableau 1).

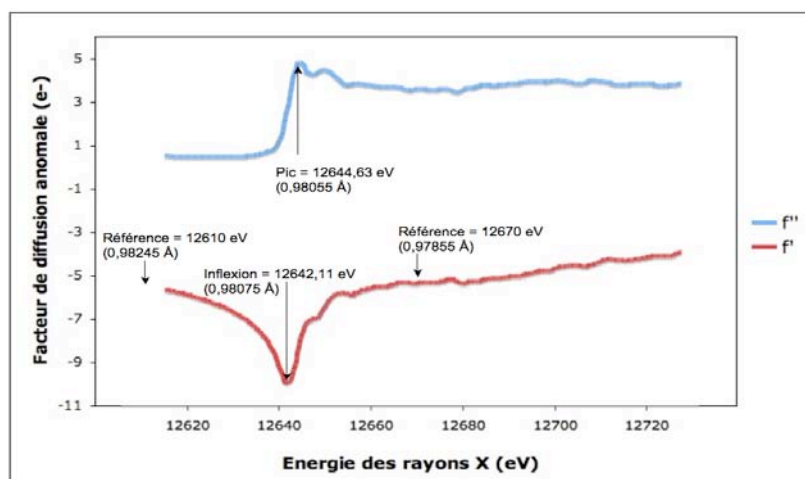


Figure 40 : Calcul des longueurs d'onde correspondant au pic d'absorption du sélénium dans la structure de SrtC-1.

II.1.4.2. Traitement des données

L'indexation et la mise à l'échelle des 4 jeux de données de SrtC-1 sélénée ainsi que du jeu de données de SrtC-1 native ont été faites avec le groupe de programmes XDS (Kabsch, 1993). Les statistiques sont rassemblées dans le tableau 1.

Collecte	SrtC-1-SeMet (pic)	SrtC-1-SeMet (inflexion)	SrtC-1-SeMet (référence1)	SrtC-1-SeMet (référence2)	SrtC-1 native
Longueur d'onde (Å)	0,98055	0,980750	0,982450	0,978550	0,93100
Groupe d'espace	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Paramètres de maille (Å)	a=68,3 b=70,1 c=86,2 α=β=γ=90°	a=68,6 b=70,4 c=86,6 α=β=γ=90°	a=68,8 b=70,5 c=86,8 α=β=γ=90°	a=68,7 b=70,3 c=86,7 α=β=γ=90°	a=68,1 b=70,2 c=86,5 α=β=γ=90°
Résolution (Å)	1,7	1,8	1,8	1,8	1,2
Nombre de réflexions uniques/totales	80448/ 297168	65317/ 240375	65393/ 240388	64343/ 236792	101238/ 1264828
Complétude (%)	96,3 (88.0)	96,7 (89.3)	95,6(84.8)	95,1(83.5)	98,3(93.9)
I/σI	13,4 (3,6)	13,7 (3,2)	12,1 (2,3)	10,1 (1,8)	33,5 (5,1)
Rsym (%)	5,9 (37,4)	5,7 (44,6)	6,1 (63,6)	7,5 (83,4)	4,7 (50,7)

Phasage	
Figure de mérite (avant/après aplatissement de solvant)	0,3/0,4

Tableau 1 : Statistiques du traitement des jeux de données de SrtC-1 native et sélénée. Les chiffres entre parenthèses représentent les valeurs de la dernière tranche de résolution.

Deux molécules sont présentes dans l'unité asymétrique. Le pourcentage de solvant est de 43% et le coefficient de Matthews de $2,17 \text{ \AA}^3$. Le phasage de SrtC-1 a été réalisé à partir des jeux de données de la protéine sélénée à une résolution de $1,73 \text{ \AA}$. Le programme AutoSHARP (Bricogne *et al.*, 2003) nous a permis de localiser les 14 méthionines sélénées de la protéine. Le programme ARP/wARP (Perrakis *et al.*, 1999), après aplatissement de solvant, a été capable de construire automatiquement 79 % de la chaîne peptidique.

Nous avons ensuite effectué un remplacement moléculaire des données natives à $1,24 \text{ \AA}$ pour améliorer la précision du modèle. L'affinement, dont les statistiques sont représentées tableau 2, a été réalisé manuellement à l'aide de COOT (Emsley and Cowtan, 2004) et du programme REFMAC5 (Collaborative Computational Project, 1994).

Affinement	SrtC-1
Nombre de résidus	390
Nombre de molécules d'eau	451
Nombre de molécules de glycérol	6
R_{work} (%)	17,1
R_{free} (%)	18,5
Facteur B moyen (\AA^2):	11,9
Déviations r.m.s pour les liaisons (\AA)	0,007
Déviations r.m.s pour les angles ($^\circ$)	1,2

Tableau 2 : Statistiques d'affinement pour SrtC-1.

II.1.5. Analyse de la structure de SrtC-1

II.1.5.1. Structure globale de SrtC-1

Nous avons pu tracer dans la carte de densité électronique les résidus 19 à 214. SrtC-1 est composée de 8 brins β centraux, principalement antiparallèles, qui s'associent pour former un tonneau, lequel est entouré de 4 hélices α majeures et de 4 hélices de type 3_{10} (figure 41). Ce type de structure en tonneau β est caractéristique du repliement des sortases de classe A et de classe B (Zhang *et al.*, 2004; Zong *et al.*, 2004a).

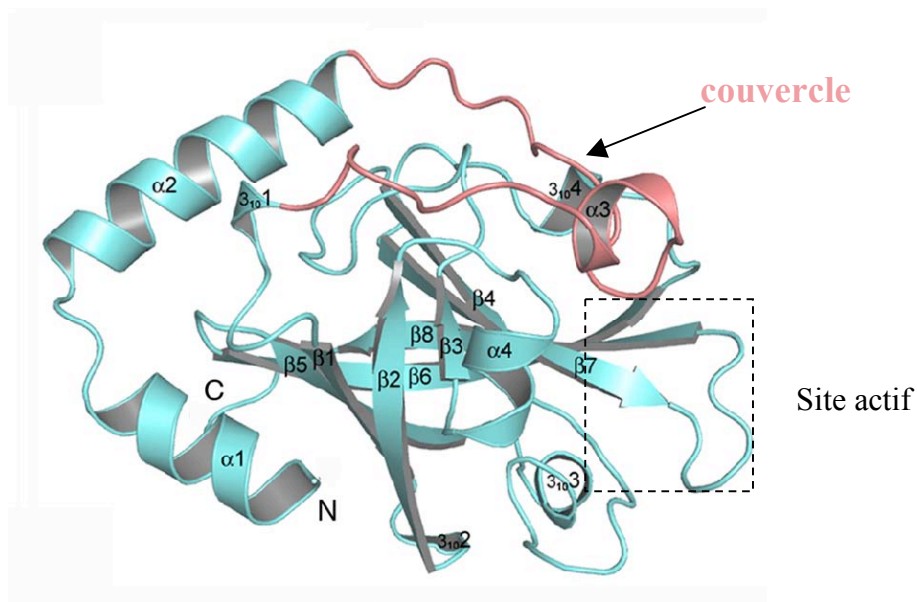


Figure 41 : Structure de SrtC-1 (Manzano *et al.*, 2008). La région centrale est composée d'un tonneau β . En rose est représenté le couvercle qui recouvre la région du site actif.

La région incluant l'hélice $\alpha 3$, colorée en rose sur la structure, forme une sorte de couvercle qui recouvre la région du site actif. Cette région semble être flexible comme en témoignent les facteurs de température qui sont plus élevés au niveau de cette région (figure 42).

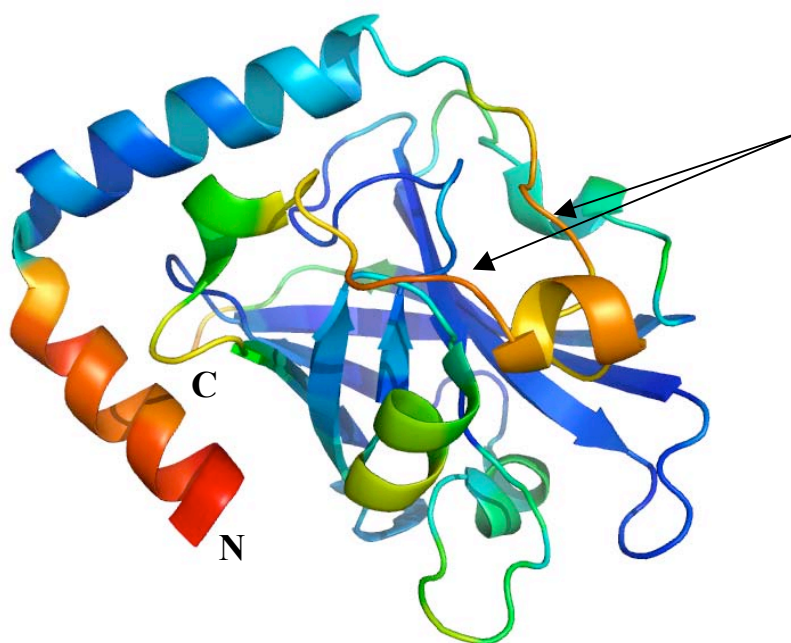


Figure 42 : Structure de SrtC-1 colorée par facteurs de température. Plus les facteurs de température sont élevés, plus la coloration est rouge. À l'inverse, plus les facteurs de température sont faibles, plus la coloration est bleue. Le couvercle semble donc être une zone potentiellement flexible comme l'indiquent les facteurs de température plus élevés des bords du couvercle (indiqués par des flèches). Il est également à noter que l'hélice en N-terminal possède des facteurs de température élevés, probablement à cause des faibles contacts qu'elle a dans l'empilement cristallin.

Alors que la région centrale de SrtC-1, englobant le tonneau β , est présente chez les sortases de classes A et B, l'arrangement des hélices α est différent. De plus, le couvercle, présent chez SrtC-1, est absent chez toutes les autres sortases dont la structure a été résolue de nos jours, aussi bien de classe A que de classe B (figure 43).

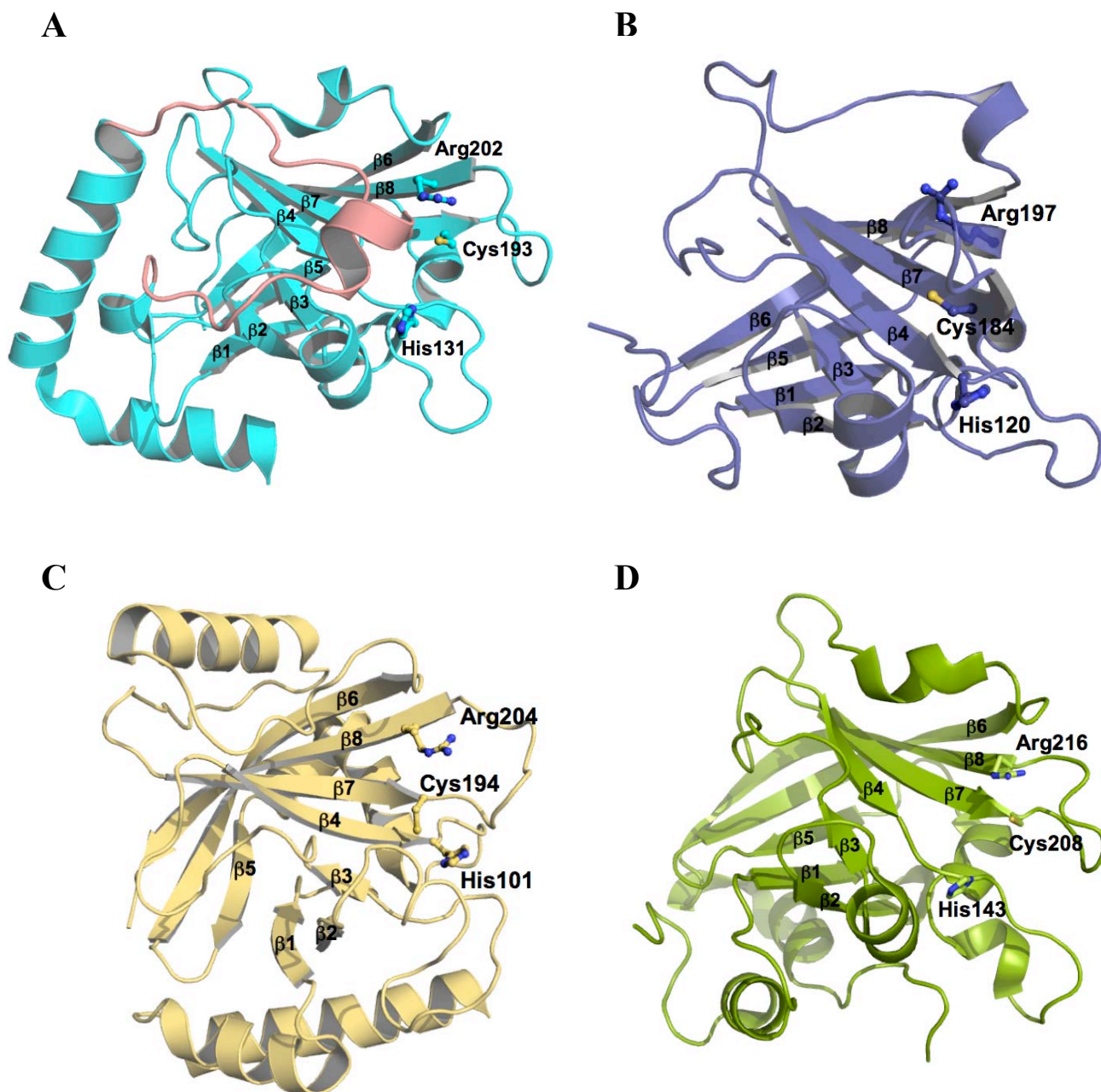
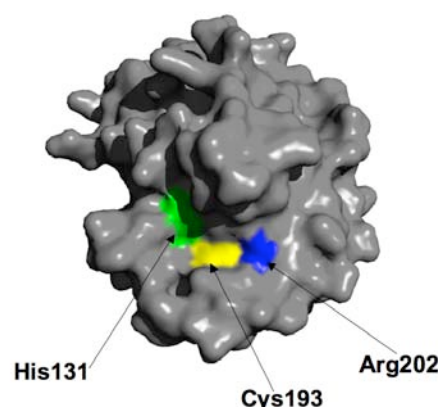


Figure 43 : SrtC-1 possède un couvercle qui recouvre la région du site actif et qui est absent chez les sortases de classe A et B. A/ Structure de SrtC-1. Le couvercle est coloré en rose. B/ Structure de Srt-A de *S. aureus*. C/ Structure de Srt-B de *S. aureus*. D/ Structure de Srt-A de *S. pyogenes*. Pour l'ensemble de ces structures, la triade catalytique, composée d'une histidine, d'une cystéine, et d'une arginine est représentée.

II.1.5.2. Site actif de SrtC-1

Le site actif des sortases de classes A et B est composé de trois résidus : une cystéine, une histidine et une arginine. SrtC-1 possède également cette triade catalytique à une position analogue de ces mêmes résidus chez les sortases de classes A et B (figure 43). On peut ainsi identifier les résidus Cys-193 (situé en C-terminal du brin β 7), His-131 (localisé dans la boucle qui suit le brin β 4) et l'Arg-202 (positionné du côté N-terminal du β 8) comme faisant partie du site actif (figure 43). Ces trois résidus sont localisés dans une gorge que l'on pourrait nommer la gorge catalytique (figure 44).

Figure 44 : Représentation en surface de SrtC-1. Les résidus du site actif (His-131, Cys-193 et Arg-202) sont localisés dans une gorge catalytique.



La chaîne latérale de Cys-193 est présente en deux conformations dans les deux monomères de l'unité asymétrique et a été représentée avec une occupation de 0,5 dans chaque cas. Dans une conformation, cette chaîne latérale pointe en dehors du site actif et fait une liaison hydrogène de 3,1 Å avec l'atome O δ 1 du résidu Asn-199, lui-même localisé sur la boucle reliant β 7 à β 8. Dans la seconde conformation, au contraire, la chaîne latérale de la Cys-193 pointe directement dans le site actif et interagit faiblement avec l'atome N de l'Arg-202 (liaison hydrogène de 3,4 Å) (figure 45A).

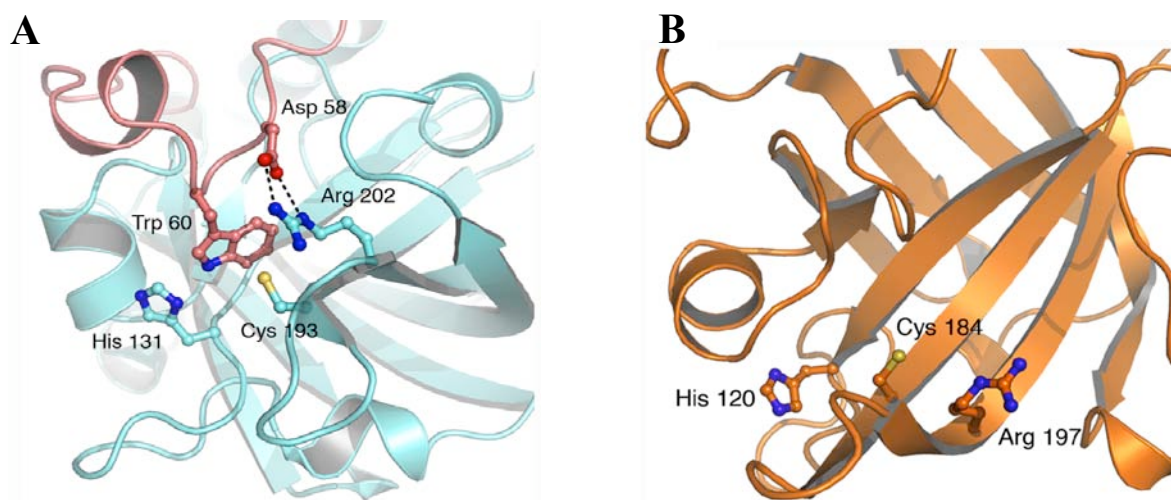


Figure 45 : Comparaison des régions du site actif de SrtC-1 de *S. pneumoniae* et SrtA de *S. aureus* (Manzano *et al.*, 2008). A/ Deux résidus, Asp58 et Trp60, permettent d'ancrer le couvercle dans le site actif de SrtC-1. B/ Le site actif de SrtA n'est pas recouvert d'un couvercle et se retrouve plus exposé au solvant.

Il est important de noter que l'His-131 n'interagit pas directement avec la Cys-193, étant donné la distance, trop importante, qu'il existe entre ces deux résidus (5,4 Å). Cette observation est en accord avec les structures d'autres sortases résolues où aucune interaction His-Cys n'a été décrite, rejetant ainsi la proposition de formation d'une paire thiolate/ion imidazolium (Zong *et al.*, 2004a; Zong *et al.*, 2004b).

Deux autres résidus, Asp-58 et Trp-60 qui appartiennent au couvercle, sont ancrés dans le site actif (figure 45A). Trp-60 augmente la nature hydrophobique de la gorge catalytique. Asp-58 interagit avec Arg-202 permettant ainsi à ce dernier résidu de pointer directement vers la cystéine catalytique. À l'inverse, la structure de SrtA de *S. aureus* montre que, d'une part, l'arginine du site actif n'est pas dans une position contrainte et que, d'autre part, elle pointe de l'autre côté de la cystéine catalytique (figure 45B).

Pour résumer, nous avons observé :

- la présence d'un couvercle, potentiellement flexible, recouvrant la région du site actif.
- l'absence de ce couvercle chez les sortases de classes A et B.
- l'ancrage du couvercle dans le site actif par le biais de deux résidus du couvercle, un tryptophane et un aspartate.
- une interaction entre l'aspartate du couvercle et l'arginine du site actif garantissant à cette dernière de pointer vers la cystéine catalytique.

L'ensemble de ces observations suggèrent que la présence d'un couvercle qui est ancré dans le site actif préviendrait la reconnaissance de substrats inappropriés. De plus, il apparaît que l'interaction entre Arg-202 et Asp-58, n'a pas seulement pour fonction de fermer le couvercle en absence de substrat, mais d'assurer aussi que Arg-202 soit dans une position optimale pour la catalyse, notamment pour stabiliser l'intermédiaire réactionnelle oxyanion (Frankel *et al.*, 2007) une fois le couvercle ouvert et le substrat lié.

II.2. Structure de SrtC-3

La production et la cristallogenèse de SrtC-3 ont été effectuées avec l'aide d'un étudiant de Master II, Thierry Izoré, que j'ai encadré.

II.2.1. Production de SrtC-3

La partie N-terminale étant prédite comme un peptide signal d'export et la partie C-terminale correspondant à un domaine transmembranaire (prédits par le logiciel SMART), seule la région incluant les acides aminés 32 à 254 de SrtC-3 a été clonée (clonage effectué par Anne-Marie Di-Guilmi). Un protocole d'expression et de purification similaire à celui de SrtC-1 a été réalisé ; SrtC-3 s'exprime avec un aussi bon rendement que celui de SrtC-1.

II.2.2. Cristallogenèse

Un criblage manuel de 168 conditions avec des kits Hampton a été effectué et a permis de révéler une condition prometteuse. Après amélioration de cette condition, un cristal est apparu en 7 jours dans 0,1 M Mes pH 5, 0,5 M $(\text{NH}_4)_2\text{SO}_4$, 1 M LiSO_4 à 20°C.

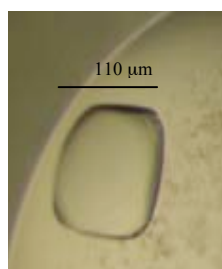


Figure 46 : Cristal de SrtC-3 dans 0,1 M Mes pH 5, 0,5 M $(\text{NH}_4)_2\text{SO}_4$, 1 M LiSO_4 .

Étant donné la haute concentration en lithium présente dans la solution de cristallogenèse, nous n'avons effectué aucune cryoprotection préalable du cristal. Le cristal a été congelé directement dans le flux d'azote liquide du générateur.

II.2.3. Résolution de la structure

Un jeu de données a été enregistré sur le générateur à rayons X du Laboratoire de Cristallogenèse et Cristallographie des Protéines de l'IBS. L'indexation et la mise à l'échelle

ont été réalisées avec le groupe de programme XDS (Kabsch, 1993). Les statistiques correspondant à ce jeu de données sont présentées dans le tableau 3.

La structure de SrtC-3 a été résolue par remplacement moléculaire à l'aide du programme PHASER (McCoy *et al.*, 2005) en prenant pour modèle la structure de SrtC-1. Le remplacement moléculaire ne fonctionnant pas avec le modèle entier de SrtC-1, nous avons choisi comme autre modèle le cœur en tonneau β de SrtC-1 qui correspond au repliement typique des sortases. Dans ce nouveau modèle (qui comprend les résidus 82 à 209), nous avons remplacé les acides aminés différents entre SrtC-1 et SrtC-3 par des alanines. La solution nous a donné 2 molécules dans l'unité asymétrique que nous avons affinées manuellement, à l'aide de COOT (Emsley and Cowtan, 2004), ainsi qu'avec le programme REFMAC5 (Murshudov *et al.*, 1997). Les paramètres d'affinement de la structure sont rassemblés dans le tableau 3.

Collecte	SrtC-3
Longueur d'onde (Å)	1,5418
Groupe d'espace	P23
Paramètres de maille (Å)	a=b=c=122,6 $\alpha=\beta=\gamma=90^\circ$
Résolution (Å)	2,14
Nombre de réflexions uniques/totales	28482/ 75646
Complétude (%)	90,7 (75,8)
I/ σ I	15,3 (4,4)
Rsym (%)	10 (51)

Affinement	
Nombre de résidus	418
Nombre de molécules d'eau	505
Nombre de molécules de MES	2
R _{work} (%)	18,1
R _{free} (%)	24,1
Facteur B moyen (Å ²):	37,68
Déviations r.m.s pour les liaisons (Å)	0,03
Déviations r.m.s pour les angles (°)	2,30

Tableau 3 : Statistiques du traitement de jeu de données de SrtC-3 et de son affinement. Les valeurs entre parenthèse sont celles de la dernière tranche de résolution.

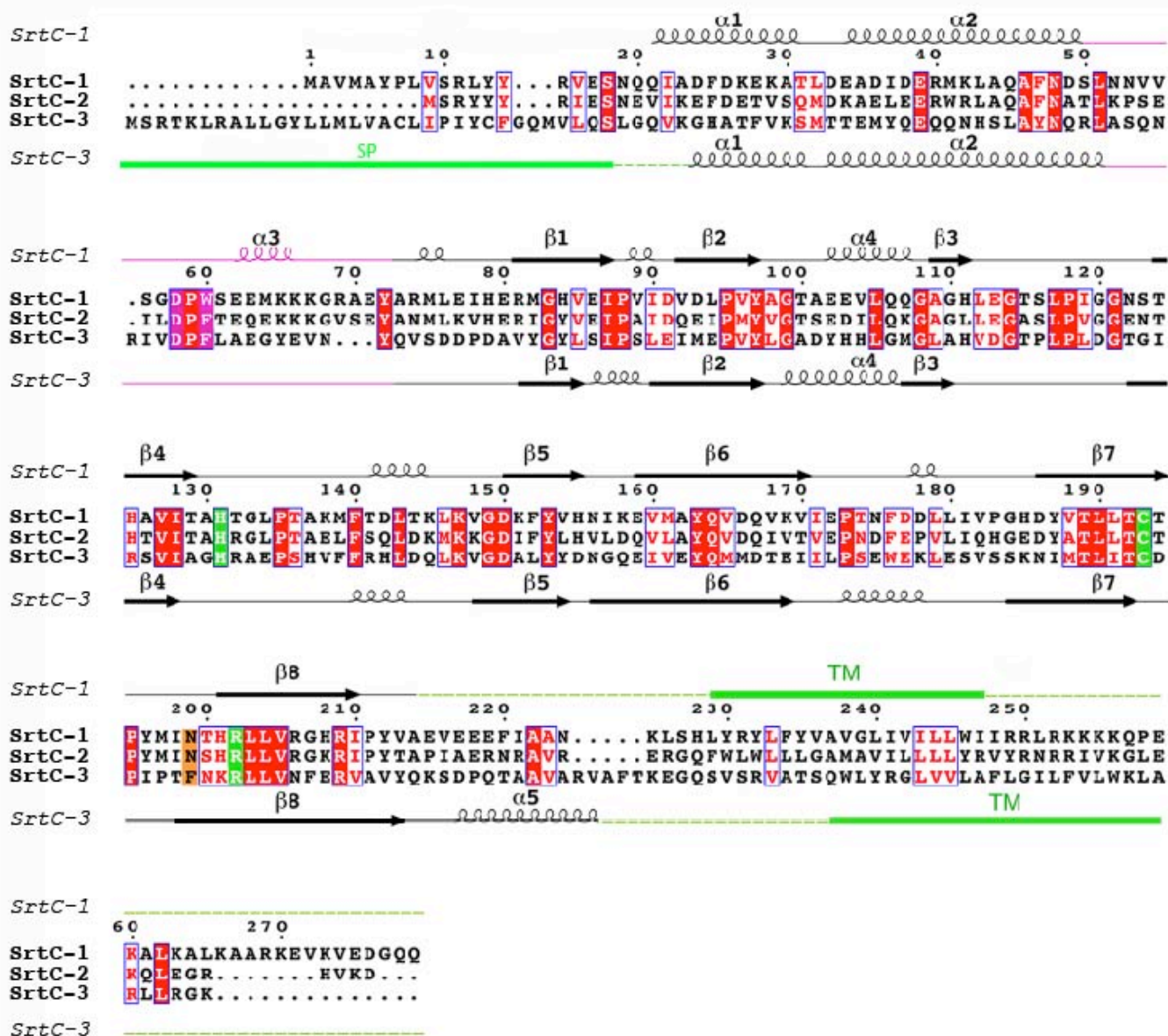


Figure 47 : Alignement de séquences des sortases de *S. pneumoniae*. Les résidus identiques sont sur fond rouge, les résidus similaires sont inscrits en rouge. Les résidues de la triade catalytique sont sur fond vert, ceux qui ancrent le couvercle dans le site actif sont sur fond rose. Le résidu Asn-199 qui interagit avec la Cys-193 est sur fond orange. La région correspond au couvercle est montrée en rose au-dessus des séquences. Les régions transmembranaires (TM) et le peptide signal (SP) ont été prédits avec le logiciel SMART (www.smart.embl-heidelberg.de). En dessus et en dessous des trois séquences se trouvent les éléments de structures secondaires de SrtC-1 et SrtC-3 respectivement.

II.2.4. Analyse de la structure de SrtC-3

II.2.4.1. Structure globale de SrtC-3

Malgré la faible identité de séquence qu'il existe entre SrtC-1 et SrtC-3 (26,4%), SrtC-3 possède un repliement très similaire à celui de SrtC-1 (se reporter à l'alignement de séquences, figure 47). Ces deux structures se superposent très bien ; la valeur du rmsd est de 1.3 Å sur 138 atomes C de la chaîne principale. Le repliement en tonneau β caractéristique aux sortases est conservé (figure 48). La principale différence structurale avec SrtC-1 réside dans la présence d'une hélice α supplémentaire du côté C-terminal.

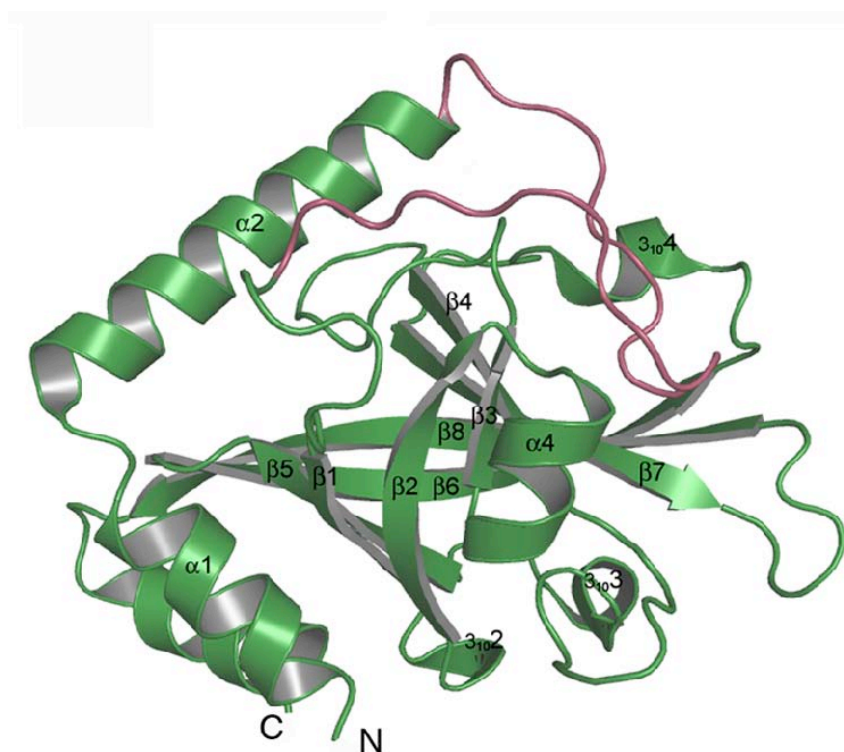


Figure 48 : Structure de SrtC-3 (Manzano *et al.*, 2008). La région du site actif recouverte d'un couvercle est montrée en rose.

Comme le montre la figure ci-dessus, SrtC-3 possède, tout comme SrtC-1, un couvercle qui recouvre la région de son site actif. Celui-ci semble être plus flexible que celui de SrtC-1, comme l'indiquent les valeurs plus élevées des facteurs de température de cette région (figure 49). Les régions comportant les résidus 68 à 71 et les résidus 78 à 82 ont, respectivement, un facteur de température de 47,4 Å² et 44,8 Å² alors que ceux de la région intérieure (résidus 73 à 76) sont plus faibles (20,2 Å²). Ces observations suggèrent que les résidus des bords du couvercle, indiqués par les flèches sur la figure 49, pourraient jouer le

rôle de « charnières » dans un potentiel mécanisme d'ouverture (telles les charnières d'un pont-levis qui permet son ouverture et sa fermeture). Cette supposition peut être extrapolée à SrtC-1 même si la différence de valeurs des facteurs de température entre la zone charnière du couvercle et la région interne ne sont pas aussi flagrantes que pour SrtC-3.

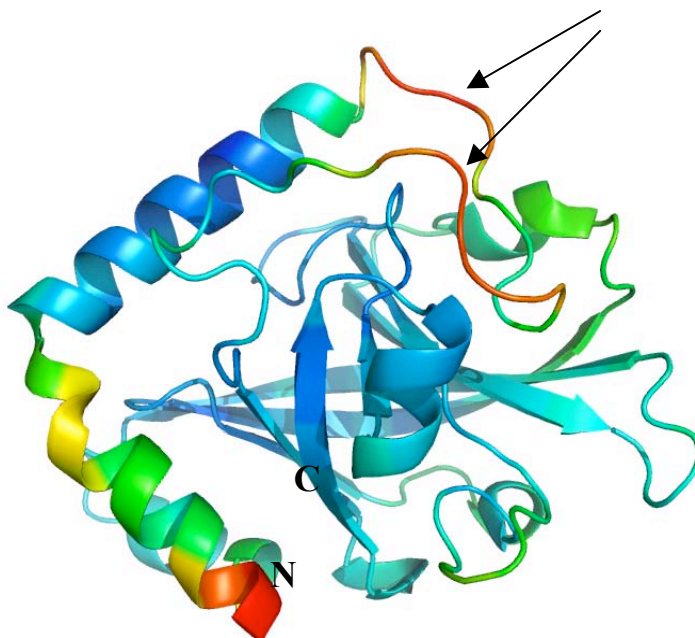


Figure 49 : Structure de SrtC-3 colorée par facteurs de température. Plus la couleur est chaude (rouge), plus le facteur de température est élevé. *A contrario*, plus la couleur est froide (bleue), plus le facteur de température est faible. Les flèches indiquent les « charnières » du couvercle.

II.2.4.2. Site actif

Le site actif de SrtC-3 montre des similarités avec celui de SrtC-1. En effet, les 3 résidus du site actif, Cys-206, His-144 et Arg-215 occupent des positions analogues aux résidus Cys-193, His-131 et Arg-202 de SrtC-1. Néanmoins, une seule conformation de Cys-206 est observée ; sa chaîne latérale pointe dans le site actif en direction de l'Arg-215 (figure 50). Dans la structure de SrtC-1, Asn-199 qui interagit avec Cys-193 expliquait la présence de la seconde conformation. Or, SrtC-3 ne possède pas une asparagine mais une phénylalanine (Phe-212), qui empêche ainsi toute interaction polaire. Ceci peut donc expliquer pourquoi une seule conformation de Cys-206 est présente dans la structure de SrtC-3. Tout comme dans SrtC-1, l'arginine du site actif de SrtC-3, Arg-215, interagit avec le résidu aspartate du couvercle (Asp-73), la maintenant de cette façon à proximité de la cystéine nucléophile. Phe-75 pointe dans la gorge catalytique lui augmentant sa nature hydrophobique (figure 50).

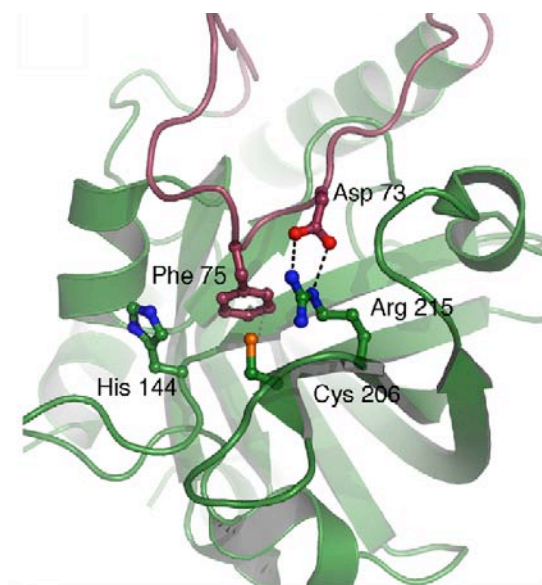


Figure 50 : Site actif de SrtC-3. Les résidus Asp73 et Phe75 appartiennent au couvercle et l'ancrent dans le site actif où les résidus de la triade catalytique sont représentés en vert.

Comme nous l'avons stipulé pour SrtC-1, SrtC-3 posséderait un couvercle permettant à l'enzyme d'être dans une conformation fermée, avec le site actif inaccessible au substrat, et garantissant que les résidus catalytiques soient positionnés de façon appropriée pour la catalyse.

II.3. La structure des sortases formant le pilus se distinguent des structures de sortases classiques

Comme décrit précédemment, SrtC-1 et SrtC-3 possèdent une région, que l'on nomme couvercle, qui recouvre leur site actif et qui est absent chez les sortases de classes A et B dont les structures ont été résolues à ce jour. Ce couvercle aurait, potentiellement, deux fonctions : celle de maintenir les résidus du site actif, notamment l'arginine, dans une position appropriée pour la catalyse et celle de restreindre l'accès du site actif afin de ne sélectionner que les substrats adéquats.

Pour étayer cette hypothèse, nous avons comparé les structures de SrtC-1 et SrtC-3 avec la structure de SrtA de *S. aureus* complexée avec le peptide mimant le motif LPXTG. Il en résulte que, de façon intéressante, le couvercle prend une position similaire à celle du peptide LPETG (figure 51). Ce peptide, dans SrtA, est localisé dans une gorge allongée, à proximité des résidus Thr-180, Leu-181, Ile-182 et Thr-183 (en jaune sur la figure 51B). Ces résidus appartiennent à la séquence consensus TLXTC, motif hautement conservé parmi

toutes les sortases et contenant la cystéine nucléophile. Il a été montré que cette région TLXTC joue un rôle dans la reconnaissance du substrat (Frankel *et al.*, 2007). Nous pouvons donc supposer que cette région, qui est aussi présente dans SrtC-1 et SrtC-3, joue un rôle dans la reconnaissance du substrat pour ces deux enzymes. Cette région, comme le montre la figure 51A, est néanmoins recouverte par une partie du couvercle et est donc inaccessible dans cet état. Une ouverture du couvercle pourrait ainsi rendre accessible cette région ainsi que le site actif. La catalyse pourrait alors se faire. Ces observations suggèrent qu'un couvercle fermé (comme observé dans les structures) prévient potentiellement la reconnaissance de la séquence LPXTG du substrat, nécessitant alors un mécanisme d'ouverture du couvercle afin que la catalyse puisse se faire.

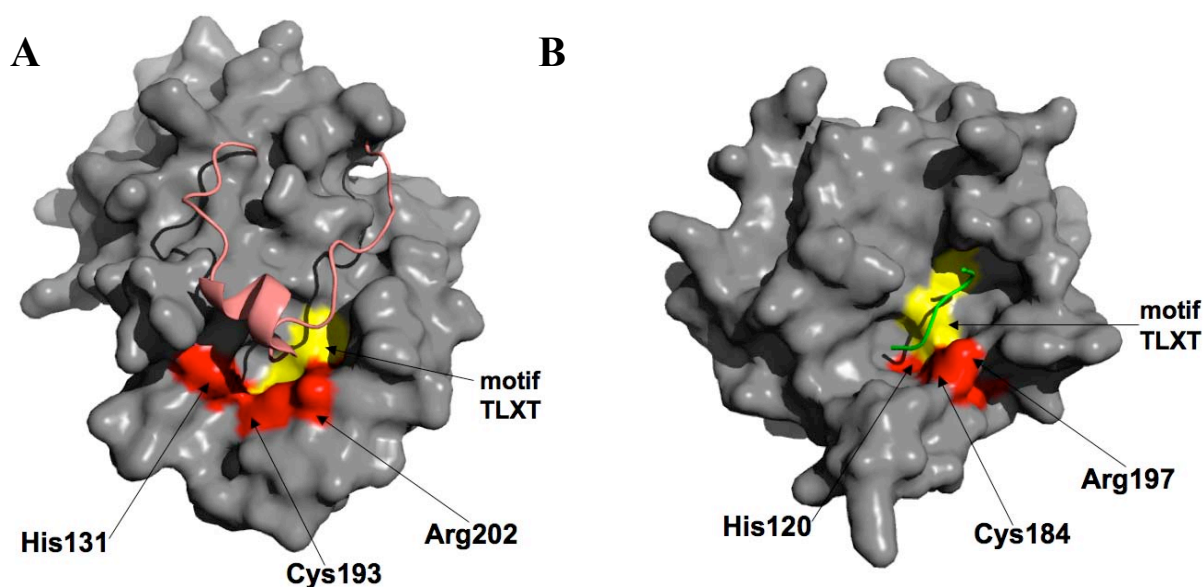


Figure 51 : Conservation de la gorge de liaison au substrat. A/ Surface de SrtC-1. Les résidus du site actif sont en rouge. La région conservée TLXT précédant la cystéine nucléophile est montrée en jaune. B/ Surface de SrtA de *S. aureus* dans laquelle les résidus du site actif et la région TLXT ont le code couleur cité précédemment. Le couvercle des sortases formant le pilus (en rose) est dans une position analogue au peptide LPETG (en vert) dans la structure de SrtA, suggérant que la gorge catalytique des sortases formant le pilus est physiquement bloquée par le couvercle.

L'ouverture du couvercle pourrait être initiée par une reconnaissance initiale entre le substrat (RrgB par exemple) et la sortase se trouvant en conformation fermée. Ainsi, seule une interaction avec le substrat approprié conduirait à l'ouverture du couvercle et à l'activation de l'enzyme. Ce modèle serait en accord avec celui proposé par Mandlik et ses collègues (Mandlik *et al.*, 2008) dans lequel ils suggèrent que pour un assemblage approprié du pilus, deux types distincts de sortases doivent interagir séquentiellement : la (ou les) sortase(s)

formant le pilus et la sortase « housekeeping ». Cette dernière reconnaîtrait le lipide II en tant que substrat et ancrerait les pili sur le peptidoglycane. Afin qu'une bonne orchestration des étapes de polymérisation et d'attachement à la paroi soit possible, il est nécessaire que les sortases formant le pilus ne reconnaissent pas le lipide II, ni les protéines substrats de SrtA. Toutefois, le site actif de ce type de sortases est assez similaire à celui des sortases de classe A, hormis la présence du couvercle recouvrant le site actif chez les sortases impliquées dans la biogenèse du pilus. De ce fait, ce couvercle, qui pourrait seulement être ouvert par une sous-unité piline appropriée, garantirait que le lipide II ne soit pas reconnu. Toutefois, Lemieux et ses collègues ont montré que SrtA de *S. pneumoniae* n'était impliquée ni dans l'assemblage et ni dans l'ancrage du pilus sur le peptidoglycane, laissant supposer que les sortases de classe C du pneumocoque seraient également capables de lier les pili à la paroi bactérienne, et donc de reconnaître le lipide II en tant que substrat. Aucune étude additionnelle concernant le rôle de SrtA n'est encore disponible et, étant donné les controverses existantes quant aux rôles joués par les différentes sortases du pneumocoque, il serait intéressant de confirmer l'action de SrtA. Il faut également soulever qu'à la surface du pneumocoque plusieurs protéines, autres que les protéines Rrg du pilus, possèdent un motif LPXTG. Ces protéines sont destinées à être ancrées sur le peptidoglycane. Une étude réalisée par Cécile Frolet du LIM montre que 15 protéines, excluant les trois protéines Rrg, chez la souche TIGR4 possèdent ce motif de reconnaissance des sortases. Ces protéines sont également présentes chez la souche R6, souche qui ne possède pas de pili. Ceci suggère ainsi que ces protéines sont ancrées au peptidoglycane par SrtA. Il est ainsi concevable de penser que ces protéines LPXTG ne doivent pas être reconnues par les sortases impliquées dans la biogenèse du pilus. La présence du couvercle sur ces dernières préviendrait ainsi la reconnaissance de ces protéines LPXTG.

En conclusion, le couvercle jouerait un rôle important pour l'activité sortasique et ne semblerait pas être simplement une région qui recouvre le site actif. De ce fait, le couvercle devrait être alors présent chez d'autres sortases de classe C. Comme aucune structure de celles-ci n'est disponible, après avoir recherché les sortases de diverses bactéries à Gram-positif ayant été étudiées et ayant un rôle dans la formation des pili, nous avons réalisé un alignement de séquences de ces différentes sortases afin de vérifier la présence du couvercle. Il semble ainsi que le couvercle, et notamment les deux résidus ancrant le couvercle dans le site actif (Asp-Trp/Phe/Tyr), soient conservés (figure 52).

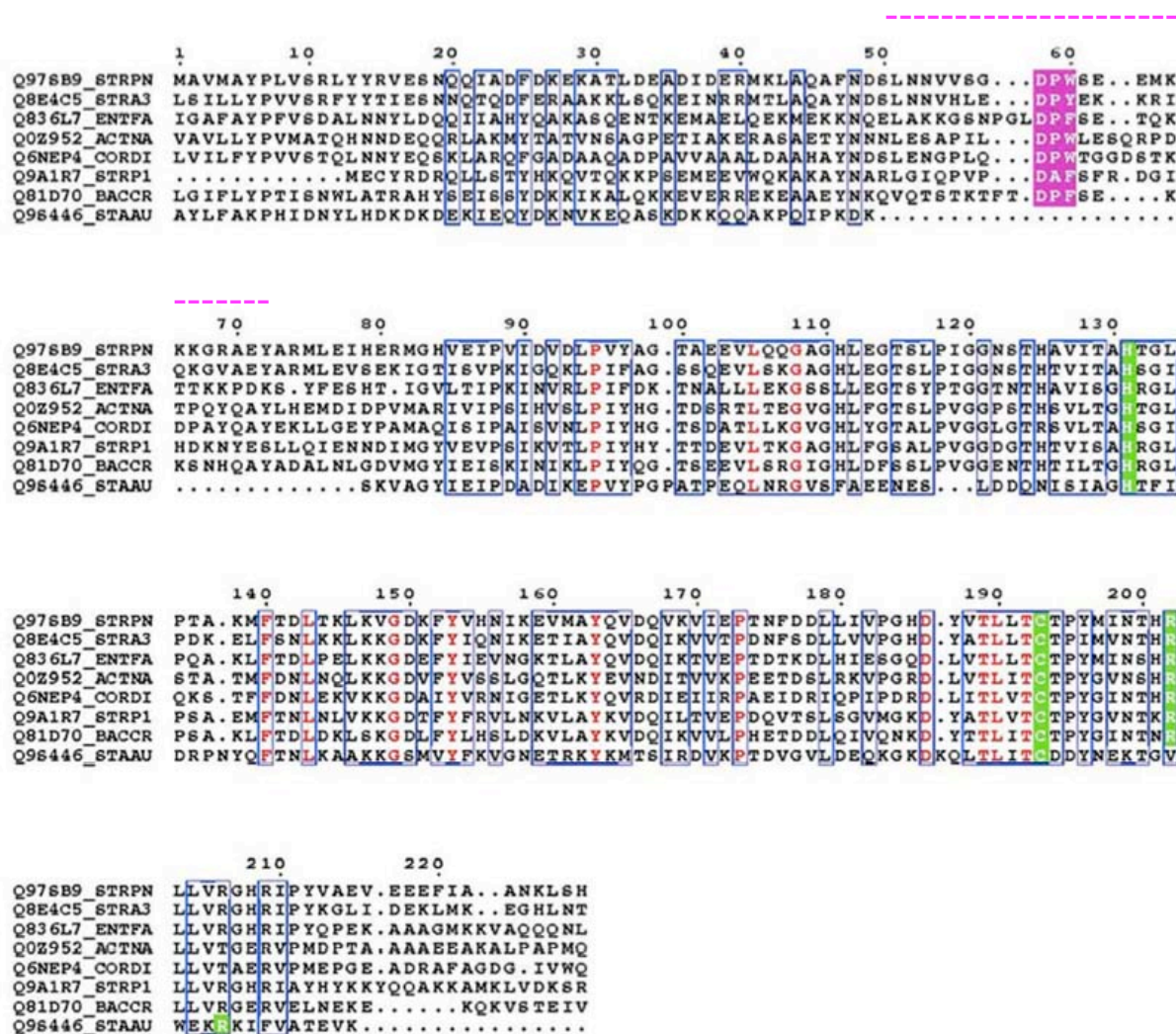


Figure 52 : Alignement de séquences de sortases formant le pilus. Séquences de *S. pneumoniae*, *S. agalactiae*, *E. faecalis*, *A. naeslundii*, *C. diphtheriae*, *S. pyogenes*, *B. cereus*, et la sortase « housekeeping » SrtA de *S. aureus* (dernière ligne). Un alignement de séquences a été effectué avec la séquence de SrtC-1. Les sortases formant le pilus possèdent un couvercle qui est absent chez les sortases non impliquées dans la biogenèse du pilus. La région correspondant au couvercle de SrtC-1 est représentée en pointillés magenta au-dessus des séquences. Les résidus conservés d'ancrage du couvercle (Asp - Pro/Ala - Trp/Tyr/Phe) sont sur fond magenta. Les résidus du site actif sont montrés en vert.

III. Caractérisation de la région du site actif de SrtC-1

Cette étude comprend l'analyse des résidus ancrant le couvercle dans la région du site actif d'une part et, d'autre part, l'analyse de la triade catalytique du site actif. Elle fait l'objet d'un second papier qui est actuellement en préparation.

III.1. Étude du couvercle de SrtC-1

La résolution de la structure de SrtC-1 (et SrtC-3) a révélé que le site actif est recouvert d'un couvercle, supposé être mobile comme l'indiquent les facteurs de température. Dans la structure de SrtC-1, ce couvercle semble être en position fermée, ancré dans le site actif par deux résidus : Asp-58 et Trp-60. Comme mentionné précédemment, nous pensons que ce couvercle constitue une région importante pour la protéine en ne restreignant, d'une part, le site actif qu'aux substrats appropriés et, d'autre part, en maintenant les résidus du site actif dans une conformation adéquate pour la catalyse. Afin d'obtenir de plus amples informations sur le rôle du couvercle, nous avons réalisé deux mutants dans lesquels les deux résidus Asp58 et Trp60 sont simultanément mutés: le double mutant SrtC-1D58GW60G et le triple mutant SrtC-1D58GW60GC193A. Nous avons décidé de faire ces mutations afin d'abolir les interactions existantes entre le couvercle et le site actif.

III.1.1. *Activité du mutant du couvercle SrtC-1D58GW60G*

Nous avons testé la capacité du double mutant SrtC-1D58GW60G de polymériser RrgB en fibre. Pour cela, comme pour les analyses *in vitro* précédemment citées, après une incubation d'une nuit à 37°C de chaque sortase (SrtC-1 sauvage et le double mutant) avec RrgB, les échantillons sont chargés sur un gel de gradient SDS. La présence de bandes de haut poids moléculaire, indiquant la présence de fibres de RrgB, est révélée par Western Blot avec les anticorps anti-RrgB. Lorsque RrgB est incubée en présence du double mutant, une faible activité de polymérisation peut être détectée, comme en témoigne le nombre restreint de bandes protéiques observées sur gel SDS (figure 53, piste 3). SrtC-1D58GW60G est donc

capable de catalyser la formation de fibres de RrgB mais moins efficacement que la protéine sauvage SrtC-1.

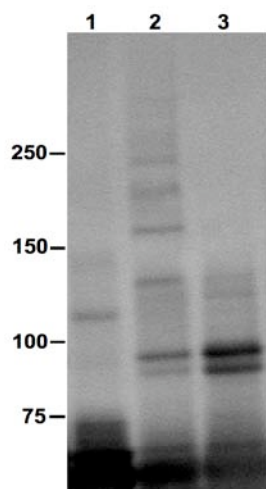


Figure 53 : Western Blot révélé avec les anticorps anti-RrgB. (1) RrgB monomérique. (2) RrgB + SrtC-1. (3) RrgB + SrtC-1D58GW60G. L'échelle de poids moléculaire est indiquée sur la gauche.

La triade catalytique (His-131, Cys-193, Arg-202) étant présente et intacte, on peut en conclure que l'abolition des interactions entre le couvercle et le site actif entraîne une diminution de l'activité catalytique ; le couvercle apparaît donc être important pour une activité réactionnelle maximale. On peut supposer que, comme l'interaction entre Asp-58 et Arg-202 n'existe plus, la conformation du résidu Arg-202 devient plus aléatoire, plus flexible et que ce résidu n'est plus maintenu dans une position optimale vers Cys-193 pour assurer son rôle dans la réaction catalytique. La même supposition peut être faite avec l'His-131 qui ne se retrouve, elle aussi, plus dans une conformation idéale pour la catalyse. Ainsi, un simple mutant, dans lequel soit Asp-58 soit Trp-60 serait muté, serait pertinent à tester. Néanmoins, la diminution de polymérisation de RrgB observée peu être imputée à une sortase qui serait moins stable donc moins efficace. Pour vérifier cela, des expériences de TSA ont été réalisées.

III.1.2. Stabilité thermique des mutants du couvercle SrtC-1D58GW60G et SrtC-1D58GW60G

Afin d'explorer l'effet des mutations glycines sur la stabilité de SrtC-1, nous avons comparé la stabilité thermique du mutant par rapport à celle du sauvage. La méthode qui a été choisie est celle de l'analyse des variations d'intensité de fluorescence en fonction de la température (Fluorescence Based-Thermal Shift Assay ou TSA). Cette technique permet de mesurer la dénaturation des protéines induite par un gradient de température et de déterminer une température de transition de dénaturation (T_m). Les résultats, présentés en figure 54,

montrent que le double mutant possède une T_m de 8°C plus basse que celle du sauvage (38°C contre 46°C). La stabilité globale de la protéine est donc affectée par les deux mutations que nous avons incorporées dans SrtC-1, pointant le fait que la chaîne latérale de Asp-58 et celle de Trp-60 sont requises pour la stabilité de la protéine. Si nous nous référons à la structure du site actif de SrtC-1, nous observons que Arg-202 est impliquée dans deux interactions avec Asp-58, résidu appartenant au couvercle. Les résultats obtenus suggèrent donc que le mutant, qui ne peut générer ces interactions, possède un couvercle qui tend à être ouvert et potentiellement flexible. Si cela est vrai, cette forme de mutant pourrait être stabilisée par l'ajout d'un substrat dans la région de son site actif.

Pour tester cette possibilité, nous avons réalisé une expérience de TSA avec le peptide synthétique IPQTG (motif LPXTG de RrgB). Afin de favoriser l'interaction entre le peptide et la protéine, nous avons travaillé avec le triple mutant SrtC-1D58GW60GC193A dans lequel la cystéine catalytique est mutée : l'interaction est ainsi soupçonnée se réaliser mais non la catalyse. La T_m du triple mutant est avant tout vérifiée et il s'avère que le triple mutant possède la même T_m que le double mutant. Le peptide IPQTG est additionné à la solution protéique à deux ratios molaires différents (peptide/protéine : 10/1 et 20 /1) avant l'analyse. Comme le montre la figure 54, il y a une augmentation de la T_m lorsque le triple mutant se trouve en présence du peptide, suggérant ainsi que ce dernier est capable d'interagir avec la protéine et de la stabiliser. De façon intéressante, plus le rapport molaire peptide/protéine augmente, plus la stabilisation se trouve renforcée.

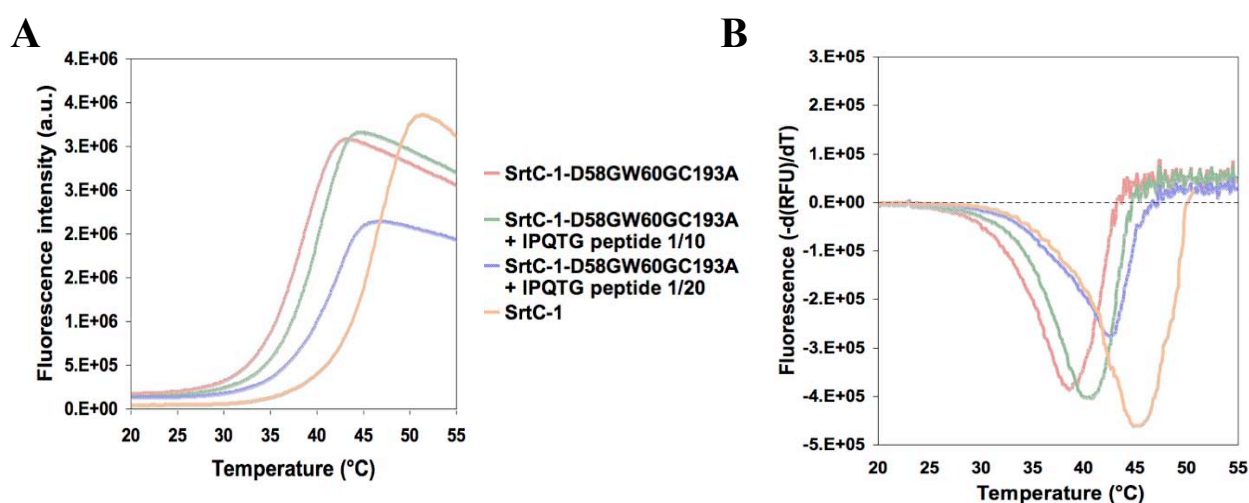


Figure 54 : Le mutant SrtC-1D58GW60GC193A est stabilisé par le peptide mimant le motif LPXTG de RrgB. Etude réalisée en Thermal Shift Assay avec 2 ratios molaires de peptides IPQTG (peptide/protéine = 10/1 et 20/1). A/ Courbes de données brutes. B/ Représentation de la dérivée première de chaque point de la courbe de fluorescence, permettant une meilleure la détermination de la T_m .

La même expérience réalisée avec la protéine sauvage SrtC-1 montre aucune amélioration de la T_m lorsque le peptide IPQTG est additionné. Le couvercle semble donc bien se retrouver ouvert lorsque les interactions entre Asp-58 et Arg-202 sont abolies, laissant ainsi le site actif accessible au ligand (ici le peptide IPQTG).

L'ensemble de ces données indiquent donc un double rôle pour la région du couvercle dans les sortases formant le pilus : (1) couvrir le site actif en absence de substrat et assurer la stabilité protéique. (2) maintenir l'Arg-202 en position optimale en direction de la Cys-193 permettant ainsi de participer à la catalyse.

III.1.3. Analyse des mutants SrtC-1_lid2 et SrtC-2_lid1

Nous avons remarqué que SrtC-2, contrairement à SrtC-1, n'était pas capable de former des fibres de RrgB *in vitro* (se reporter à la figure 32). Afin d'étudier le rôle du couvercle et son implication dans la reconnaissance du substrat, nous avons inversé le couvercle de ces deux sortases et étudié si elle était toujours capable de polymériser RrgB. Deux mutants ont été ainsi construits : le mutant SrtC-1 qui possède le couvercle de SrtC-2 (appelé SrtC-1_lid2) et le mutant SrtC-2 ayant le couvercle de SrtC-1 (appelé SrtC-2_lid1). Ces mutants ont été testés dans leur capacité à polymériser RrgB en fibre *in vitro*. Contrairement à SrtC-1, le mutant SrtC-1_lid2 ne polymérise pas RrgB. Tout comme SrtC-2, le mutant SrtC-2_lid1 est incapable de former des fibres de RrgB. Ces résultats peuvent ainsi suggérer que le couvercle est une zone importante pour la reconnaissance du substrat (puisque RrgB n'est plus reconnu par SrtC-1_lid2 comme substrat) mais pas exclusivement. En effet, la surface de SrtC-1 doit également être importante pour la reconnaissance du substrat car SrtC-2_lid1 n'est toujours pas capable de polymériser RrgB.

Ces résultats restent néanmoins préliminaires et restent à reproduire. De plus, il serait intéressant de comparer la stabilité des deux mutants à celle des deux sortases sauvages. En effet, on peut penser qu'une perte de l'activité du mutant SrtC-1_lid2 est la conséquence d'une protéine moins stable.

III.2. Étude du site actif de SrtC-1

Les sortases « housekeeping » catalysent l'attachement de protéines cibles à la surface de la bactérie grâce à un mécanisme qui fait intervenir 3 résidus conservés et essentiels dans toutes les sortases: une Cystéine, une Histidine et une Arginine. Il est ainsi question de triade catalytique pour désigner le site actif.

Aucune information concernant un éventuel rôle de l'arginine et de l'histidine du site actif des sortases impliquées dans la formation du pilus n'est disponible à ce jour, pouvant ainsi émettre un doute sur la conservation ou non de ces résidus dans le mécanisme catalytique. La résolution de la structure de SrtC-1 a révélé que His-131 et Arg-202 étaient ancrées dans le site actif à des positions analogues à celles de His-120 et Arg-197 de SrtA de *S. aureus*, résidus clé du mécanisme catalytique. Afin de vérifier si ces résidus His-131 et Arg-202 sont importants pour l'activité de SrtC-1 (et donc la formation du pilus), deux mutants du site actif ont été construits : SrtC-1H131D et SrtC-1R202E. Nous avons décidé de muter l'histidine en aspartate et l'arginine en glutamate afin de modifier la charge tout en gardant une longueur à peu près analogue de la chaîne latérale. Ces deux mutants ont été étudiés dans leur capacité à polymériser RrgB en fibre afin de déterminer si les résidus His-131 et Arg-202 jouent un rôle dans l'activité enzymatique de SrtC-1. Une étude de leur stabilité thermique a été également effectuée. De plus, la structure d'un des deux mutants a pu être obtenue par cristallographie aux rayons X.

III.2.1. *Activité des mutants du site actif*

Après expression et purification simultanée de SrtC-1, SrtC-1H131D, SrtC-1R202E et RrgB, la capacité de chaque mutant à polymériser le monomère RrgB est analysée par le même protocole qu'avec le mutant du couvercle.

Aucune bande de haut poids moléculaire, relative à une éventuelle polymérisation de RrgB, n'est détectée lorsque RrgB est mise en présence de l'un des deux mutants SrtC-1H131D ou SrtC-1R202E (figure 55, pistes 3 et 4). Cela suggère que les résidus His-131 et Arg-202 jouent un rôle clé dans le processus de polymérisation du pilus. Il s'avère, par conséquent, que SrtC-1, tout comme les sortases de classes A et B, possède bien une triade catalytique composée de Cys-193, His-131 et Arg-202. Les résidus impliqués dans le

mécanisme catalytique des sortases de classe C semble donc être les mêmes que ceux des sortases impliquées dans l'ancrage de protéines cibles sur le peptidoglycane.

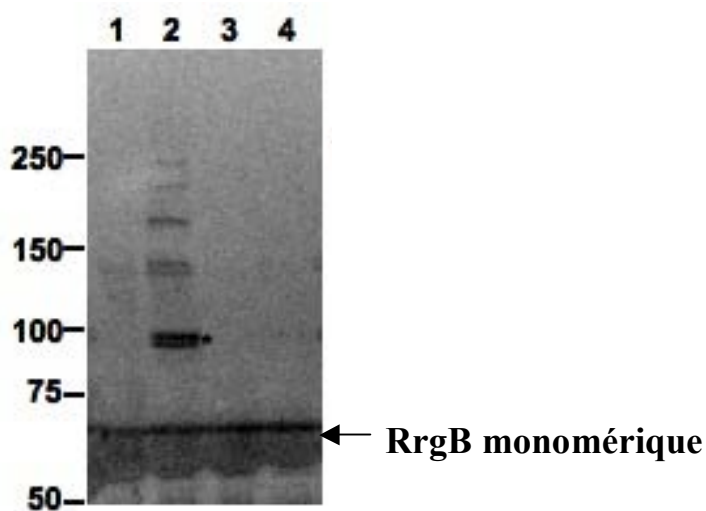


Figure 55 : Western Blot révélé avec des anticorps dirigés contre RrgB. (1) RrgB monomérique. (2) RrgB + SrtC1. (3) RrgB + SrtC-1H131D. (4) RrgB + SrtC-1R202E. L'échelle de poids moléculaire est montrée sur la gauche.

Si nous réalisons un alignement de séquences de plusieurs sortases de divers organismes impliquées dans la formation de pilus, nous pouvons également observer que les résidus Cystéine, Histidine et Arginine sont conservés, suggérant ainsi leur conservation dans le mécanisme catalytique chez les bactéries à Gram-positif (figure 52).

III.2.2. Stabilité thermique des mutants

Afin d'étudier si les mutations que nous avons introduites engendraient des perturbations conformationnelles, des expériences de TSA ont été initiées. Les profils de dénaturation de SrtC-1, SrtC-1H131D et SrtC-1R202E, montrés ci-dessous, révèlent que SrtC-1 et le mutant SrtC-1H131D possèdent une T_m d'environ 46°C alors que la T_m du mutant SrtC-1R202E est plus faible (environ 39°C). La mutation de l'histidine en aspartate ne semble donc pas affecter la stabilité thermique (et donc la structure globale) de la protéine alors que la mutation de l'arginine en glutamate conduit à une déstabilisation conformationnelle de la protéine. Le mutant SrtC-1R202E se conduit de la même façon que le double mutant SrtC-1D58GW60 en TSA, mais non en activité puisqu'il est totalement inactif. Ces observations

suggèrent ainsi, que la perte d'activité du mutant SrtC-1R202E est la conséquence de la perte du résidu catalytique Arg-202 et non de l'instabilité de la protéine.

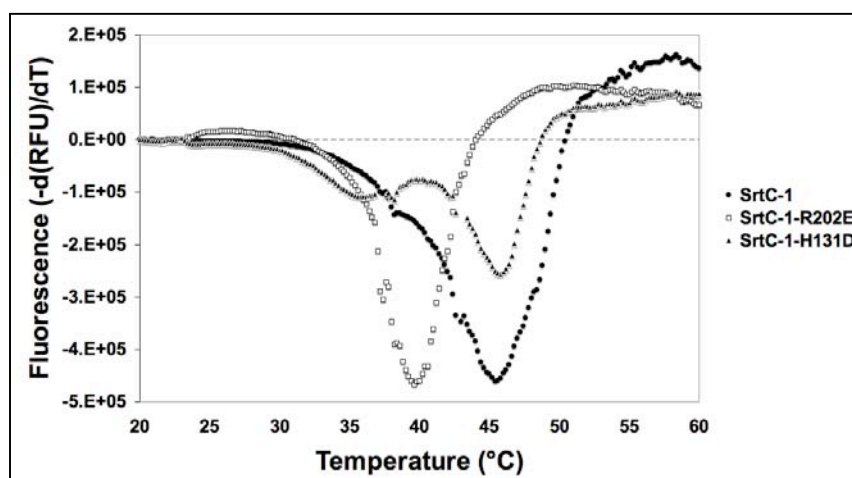


Figure 56 : Stabilité thermique de SrtC-1 et des mutants du site actif évaluée par Thermal Shift Assay. La dérivée première de la fluorescence permet la détermination de la T_m .

Arg-202 interagit avec Asp-58, résidu appartenant au couvercle. Il apparaît ainsi que, au moins en absence de substrat, l'interaction entre Asp-58 et Arg-202 est essentielle pour le couvercle et la stabilité globale de la protéine. Il est ainsi tentant de supposer que lorsqu'on supprime ces interactions, le couvercle n'est plus maintenu dans une conformation fermée. La conformation de la protéine ayant son couvercle ouvert serait alors moins stable, ce qui expliquerait les observations obtenues en TSA. En outre, aucun cristal de SrtC-1R202E n'a pu être obtenu confortant l'idée de la présence d'un couvercle trop flexible pour permettre la cristallisation. Le mutant SrtC-1H131D, quant à lui, est cristallisé dans 0,1 M Tris pH 8,5, 0,2 M $MgCl_2$ hexahydrate, 30 % PEG4000. Afin d'obtenir une vue du site actif au niveau atomique, la structure de SrtC-1H131D a été déterminée à haute résolution (1,3 Å) par cristallographie aux rayons X. Cette structure a été obtenue par remplacement moléculaire avec la structure native. SrtC-1H131D est dotée d'un repliement similaire à celui de la protéine native (rms deviation = 0,110 Å sur 184 atomes). L'analyse du site actif révèle que Asp-131 interagit avec l'atome N du Trp-60 (figure 57). Le couvercle se trouve toujours dans la même position, ce qui peut expliquer pourquoi la stabilité thermique de la protéine n'est pas affectée.

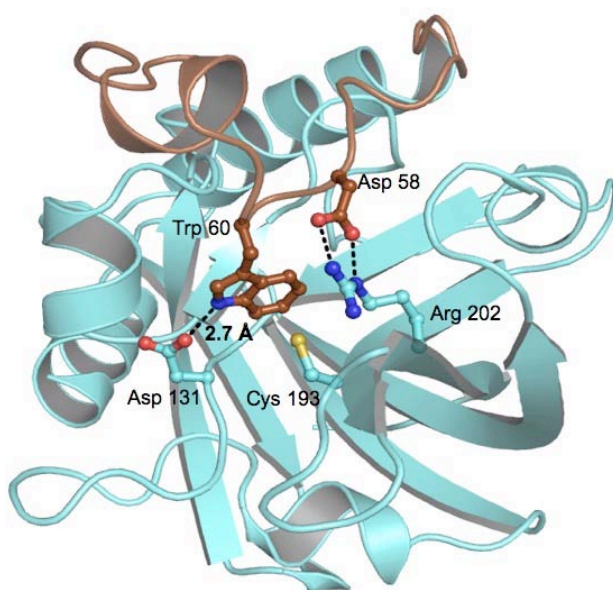


Figure 57 : Structure du site actif du mutant SrtC-1His131Asp. Le résidu muté, Asp-131 interagit avec Trp-60.

L'ensemble de ces études suggèrent donc que la formation des fibres de RrgB par SrtC-1 requière une triade catalytique composé de His-131, Cys-193 et Arg-202.

IV. Formation *in vitro* d'un complexe covalent entre SrtC-1 et RrgB

Il était, certes, cohérent de révéler les Westerns-Blot avec les anticorps dirigés contre RrgB pour analyser la formation de fibres, mais il était également intéressant d'utiliser les anticorps dirigés contre SrtC-1. Premièrement, la présence d'une seconde bande protéique migrant aux alentours de 60 kDa est allumée avec les anticorps anti-SrtC-1 après que SrtC-1, seule ou en présence de RrgB, soit incubée une nuit à 37°C (figure 58). Cette observation suggère que SrtC-1 a tendance à s'oligomériser. Il est supposé que ce soit le dimère que nous voyons sur gel étant donné que SrtC-1 monomérique migre aux alentours de 30 kDa. Il a été montré que SrtA de *S. aureus*, en solution, s'oligomériser et peut se trouver sous forme monomère, dimère, trimère ou encore quadrimère, les deux dernières formes étant largement minoritaires (Lu *et al.*, 2007). Il n'est donc peut-être pas si surprenant de voir SrtC-1 dimériser également. D'autre part, de façon très intéressante, une bande migrant vers 85 kDa (indiquée par une étoile sur le gel de la figure 58) est révélée à la fois avec les anticorps anti-SrtC-1 et anti-RrgB lorsque RrgB est incubée avec SrtC-1 à 37°C durant une nuit. Ceci pourrait indiquer la présence d'un complexe covalent (car non dissocié par chauffage et par le SDS) entre SrtC-1 et RrgB.

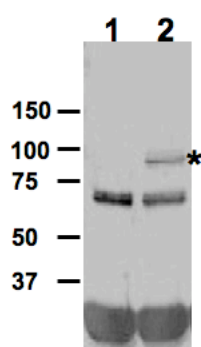


Figure 58 : A/ Western Blot révélé avec les anticorps dirigés contre SrtC-1. (1) SrtC-1. (2) SrtC-1 + RrgB. L'étoile révèle la bande à la fois allumée par les anticorps anti-SrtC-1 et anti-RrgB.

Afin de confirmer cela, des analyses de masse ont été réalisées. Tout d'abord, l'échantillon SrtC-1+RrgB, après incubation à 37°C pendant une nuit, est analysé en spectrométrie de masse MALDI-TOF au Laboratoire de Spectrométrie de Masse des Protéines de l'IBS. Une espèce protéique de 88,3 kDa est détectée, poids moléculaire correspondant à un complexe 1/1 SrtC-1/RrgB (Figure 59A). Une masse correspondant à un dimère RrgB est

également observée, cependant, les espèces de poids moléculaires supérieurs ne sont pas détectées en raison de leur faible concentration dans l'échantillon. Une analyse de spectrométrie de masse en tandem couplée à une chromatographie en phase liquide (nanoLC-MS/MS) est de plus effectuée au CEA grâce à la plateforme « EDyP-Service ». Pour ce faire, la bande supposée être le complexe SrtC-1/RrgB est découpée du gel SDS, soumis à une digestion trypsique et les fragments générés, après avoir été séparés par chromatographie d'exclusion de taille, sont analysés. Deux séquences qui correspondent à SrtC-1 et RrgB sont retrouvées dans cette bande (figure 59B). La bande présente sur gel SDS, allumée à la fois par les anticorps anti-SrtC-1 et anti-RrgB, contient donc bien un complexe covalent SrtC-1/RrgB.

A

Masse Moléculaire expérimentale (Da)	Masse moléculaire calculée (Da)	Espèce protéique correspondante
23786	23782.9	SrtC-1
64933	65054.5	RrgB
88372	88837.4	Complexe SrtC-1/RrgB

B

Masse observée	Masse expérimentale	Masse calculée	Δ	Premier AA	Dernier AA	Séquence	Protéine identifiée
911.57	1821.13	1820.91	0.21	210	225	IPYVAEVEEFIAANK	SrtC-1
663.35	1324.70	1324.63	0.07	159	169	EVMAQVDQVK	SrtC-1
508.28	1521.82	1521.69	0.13	30	42	ATLDEADIDERMK	SrtC-1
523.84	1045.67	1045.56	0.11	129	138	FNTANLPAAK	RrgB
563.35	1124.69	1124.57	0.12	300	309	ITYSATLNDK	RrgB
588.87	1175.72	1175.63	0.09	580	590	QPAGYALLTSR	RrgB
602.88	1203.75	1203.67	0.08	377	387	VVQTVTLTTDK	RrgB
740.43	1478.85	1478.72	0.12	431	443	DENPKPLDPEPK	RrgB
566.01	1695.02	1694.85	0.17	388	402	NTVTVNGLDKNTYK	RrgB
639.39	1915.14	1914.92	0.22	410	427	GYSADYQEITTAGEIAVK	RrgB
688.76	2063.25	2063.04	0.21	461	480	LAGAEFVIANADNAGQYLAR	RrgB
749.49	2245.45	2245.22	0.24	377	397	VVQTVTLTTDKNTVTVNGLDK	RrgB
781.14	2340.39	2340.15	0.24	141	162	IYEIHSLSYVGEDGATLTGSK	RrgB
828.85	2483.54	2483.25	0.28	206	227	VDKDPVNHQVGDVVEYEVTK	RrgB
857.17	2568.49	2568.20	0.29	42	65	LLATDGDMDKIANELETGNYAGNK	RrgB

Figure 59 : Etude du complexe covalent SrtC-1/RrgB par spectrométrie de masse. A/ Analyse du mélange SrtC-1 + RrgB après un nuit d'incubation à 37°C par spectrométrie de masse MALDI. B/ Analyse par nanoLC-MS/MS de la bande sur gel correspondant au complexe SrtC-1/ RrgB. AA signifie acide aminé.

En outre, ce complexe covalent peut être détecté avec le mutant SrtC-1D58GW60G mais non avec les mutants SrtC-1H131D et SrtC-1R202E. Ces résultats indiquent que la polymérisation de RrgB se réalise à travers la formation d'un intermédiaire acyl-enzyme dont l'élaboration est dépendante d'une triade catalytique intacte. Le fait que le complexe SrtC-1/RrgB peut être observé avec le mutant SrtC-1D58GW60G indique que la mutation des résidus ancrant le couvercle ne bloque ni la reconnaissance du substrat, ni la formation d'un intermédiaire acyl-enzyme mais affecte la procession de la réaction. En revanche, l'absence de complexe SrtC-1/RrgB et de fibres détectées avec les mutants SrtC-1H131D et SrtC-1R202E suggèrent que les deux résidus His-131 et Arg-202 participent à une étape réactionnelle qui a lieu avant la formation d'un intermédiaire acyl-enzyme. Dans SrtA de *S. aureus*, le résidu His est supposé jouer un rôle dans la protonation du groupe partant et le résidu Arg stabiliserait la charge négative de l'état de transition (l'oxyanion) ; ces deux événements facilitant ainsi la formation d'un intermédiaire acyl-enzyme (Frankel *et al.*, 2005 ; Frankel *et al.*, 2007). Nos résultats paraissent donc en accord avec cela.

Il serait intéressant d'isoler ce complexe dans le but d'effectuer une étude structurale par cristallographie aux rayons X. Dans ce cas, une incubation à large échelle de SrtC-1 avec RrgB pourrait être faite à 37°C pendant une nuit suivie d'une gel filtration afin d'éliminer les monomères de SrtC-1 et de RrgB. Une autre alternative pourrait être envisagée afin de bien séparer le complexe du monomère RrgB (en effet, en gel filtration, le complexe est toujours contaminé par du RrgB monomérique. Ainsi, la protéine SrtC-1 étiquetée pourrait être utilisée pour réaliser l'incubation à 37°C sur la nuit avec RrgB, le mélange pourrait être chargé ensuite sur colonne d'affinité HisTrapTMHP. Après élution, les protéines éluées seraient chargées sur une gel filtration afin de séparer le monomère SrtC-1 du complexe SrtC-1/RrgB. Des cribles à l'aide du robot de cristallisation disponible à l'EMBL pourraient être ensuite effectués. L'obtention de la structure d'un tel complexe permettrait ainsi de comprendre un peu mieux le mécanisme catalytique des sortases impliquées dans la formation du pilus.

V. Étude structurale de RrgB

Des études de cristallogénèse avec la protéine RrgB ont été initiées. Des cribles effectués de façon manuelle ont été réalisés avec les kits Hampton. Une piste prometteuse a été obtenue dans 0,1 M Hepes pH 7,5, 20% PEG10000 avec une concentration protéique de 10 mg/ml. Cette condition a été améliorée et des cristaux ont été obtenus dans 0,1 M Tris pH 7,5, 18% PEG12000 (figure 60).



Figure 60 : Cristaux de RrgB obtenus dans 0,1 M Tris pH 7,5 et 18% PEG12000.

Ces cristaux poussent en un peu moins de deux mois. Après cryoprotection des cristaux dans une solution-mère supplémentée en 20% glycérol, ceux-ci ont été refroidis dans l'azote liquide. Un jeu de données à 3,7 Å a été enregistré sur la ligne de lumière ID14-EH1 de l'ESRF, avec un pas d'oscillation de 0,5°, pour éviter le recouvrement des taches (figure 61). Ce jeu de données a été traité avec le groupe de programmes XDS. Les statistiques de ce jeu sont décrites dans le tableau.

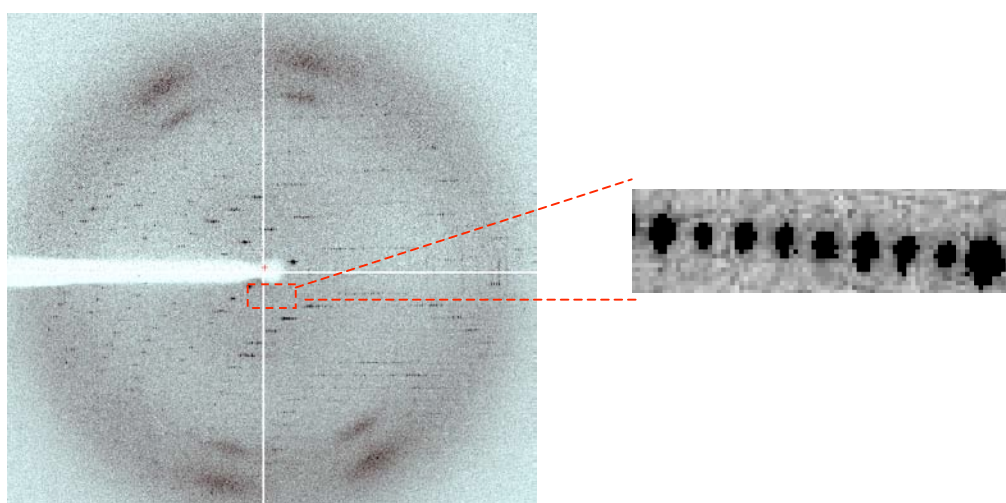


Figure 61 : Clichés de diffraction d'un cristal de RrgB native diffractant à 3.7 Å pris à 90°. Un zoom est effectué sur la droite permettant de voir que les taches de diffraction sont assez rapprochées mais séparées les unes des autres.

Après purification de RrgB en milieu sélénié, des cristaux de RrgB séléniée ont été obtenus dans 0,1 M Tris pH 7,5, 23% PEG8000. Après une cryoprotection similaire à celle des cristaux de RrgB native, un jeu de données a été collecté à 3 Å avec un pas d'oscillation de 0,1°, permettant ainsi de bien séparer les taches de diffraction (figure 62).

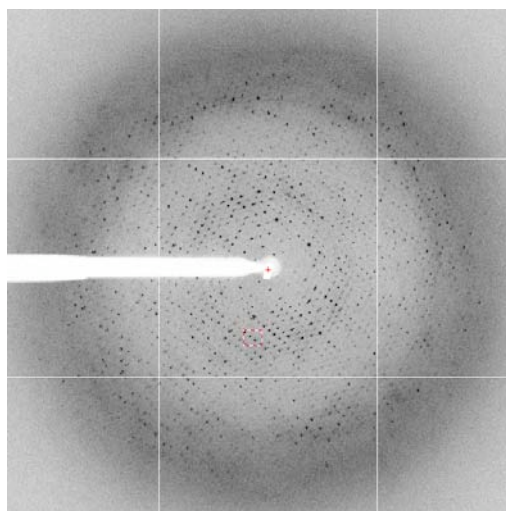


Figure 62 : Cliché de diffraction d'un cristal de RrgB séléniée diffractant à 3 Å, pris à 90°.

Nous avons collecté un jeu de données au seuil du sélénium à la longueur d'onde correspondant au pic. Ce jeu a été indexé et mis à l'échelle grâce au groupe de programme XDS. Les paramètres de maille du cristal de RrgB séléniée sont différents de ceux du cristal de RrgB natif, 12 molécules par unité asymétrique sont ainsi présentes (se reporter au tableau). Nous avons tenté d'obtenir les phases par SAD. Malheureusement, le signal anomal était trop pauvre pour pouvoir résoudre la structure par SAD. Nous projetons de tester d'autres cristaux de RrgB native afin d'une part d'améliorer la résolution et, d'autre part, de trouver des paramètres de maille similaire à ceux du cristal de RrgB séléniée. Nous envisageons également de refaire des cristaux de RrgB séléniée afin de collecter un jeu de données avec des paramètres de maille similaires à ceux du cristal de RrgB native, où seules 3 molécules par unité asymétrique sont présentes.

Collecte	RrgB native	RrgB SeMet (pic)
Longueur d'onde (Å)	0,933400	0,97950
Groupe d'espace	P321	P3 ₂
Paramètres de maille (Å)	a=72,5 b=72,5 c=340,1 $\alpha=\beta=90^\circ$ $\gamma=120^\circ$	a=140,3 b=140,3 c=341,2 $\alpha=\beta=90^\circ$ $\gamma=120^\circ$
Résolution (Å)	3,7	3,1
Nombre de réflexions uniques/totales	11675/243901	276029/ 767713
Complétude (%)	97,3 (91,6)	90,7 (86,8)
I/ σ I	17,9 (8,0)	7,2 (2,8)
Rsym (%)	21,2 (51,4)	12,2 (42,0)

Tableau 4 : Statistiques du traitement des jeux de données de RrgB native et sélénée. Les valeurs entre parenthèse sont celles de la dernière tranche de résolution.

Conclusions et Perspectives

Quelques bactéries à Gram-positif, dont certaines souches sont virulentes pour l'homme, ont développé des systèmes de pathogénie, les pili, qui leur permettent d'adhérer aux cellules hôtes. Ces pili, qui sont des structures protéiques filamenteuses, sont assemblés par des enzymes spéciales, les sortases. Chez *S. pneumoniae*, la présence de pili est corrélée avec la présence d'un îlot de pathogénicité *rlrA*. Celui-ci se compose de trois gènes codant pour des protéines qui structurent le pilus (RrgA, RrgB et RrgC) mais également de trois gènes codant pour des sortases de classe C (SrtC-1, SrtC-2 et SrtC-3).

Rôle des sortases dans la polymérisation de RrgB.

Découvrir le ou les acteur(s) sortasique(s) formant le corps du pilus est une question essentielle sur laquelle nous nous sommes penchés. En effet, des travaux antérieurs ont montré que sans polymérisation de RrgB, les souches pneumococciques sont déficientes en pili et sont, dès lors, moins adhérentes sur les cellules hôtes et moins virulentes. Ainsi, le rôle que joue la (ou les) sortase(s) qui polymérise RrgB en fait d'elle(s) une cible attrayante pour le développement d'inhibiteurs. C'est une approche biochimique, en travaillant avec les protéines recombinantes, que nous avons utilisé afin de déterminer l'implication des trois sortases, SrtC-1, SrtC-2 et SrtC-3, dans la polymérisation de RrgB en pilus.

De nombreux efforts ont été fournis en terme de purification de protéines et de tests pour mettre au point le système de fabrication de fibres *in vitro* et démontrer ainsi le rôle de chacune des trois sortases dans la polymérisation de RrgB. Il a été ainsi montré que SrtC-1 est la principale sortase impliquée dans la polymérisation de RrgB en fibre bien que SrtC-3 montre également une faible activité de polymérisation. C'est la première fois que l'activité de sortases de classe C a pu être mise en évidence d'une manière réalisée *in vitro*. En effet, seules des études réalisées *in vivo*, sur des souches délétées de gènes de sortases, étaient disponibles. Ainsi, disposer d'un test reproductible, facile à mettre en œuvre, peut constituer un atout dans des essais d'inhibition d'activité sortasique. Nous avons été également capables d'observer les fibres formées *in vitro* par microscopie électronique. Ceci nous a permis de mieux caractériser les fibres et de réaliser que les « objets » formés *in vitro* avaient un sens : nos fibres sont flexibles, d'un diamètre de 3,5 à 7 nm ce qui corrèle avec les mesures prises *in vivo*.

Notre étude montre que SrtC-1 est la sortase principale impliquée dans la formation de RrgB en fibres et qu'elle pourrait être ainsi une cible thérapeutique potentielle. Il est toutefois important de noter que l'activité de formation du corps de pilus peut être réalisée

potentiellement par une autre sortase : dans nos travaux, nous montrons que, même en absence de SrtC-1, la présence de pili est retrouvée à la surface pneumococcique, ce qui est également confirmé par les travaux de LeMieux et ses collègues ainsi que de Falker et ses collègues. De ce fait, cibler les trois sortases formant le pilus se révélerait être plus pertinent pour inhiber la polymérisation du pilus.

Les sortases impliquées dans la formation du pilus chez *S. pneumoniae* possèdent un couvercle qui est absent chez les sortases de classes A et B.

Des essais de cristallogénèse ont été initiés avec les trois sortases et deux structures de sortases formant le pilus, SrtC-1 et SrtC-3, ont été pour la première fois résolues. Ces structures révèlent une organisation similaire aux sortases impliquées dans l'ancrage de protéine au peptidoglycane. Elles possèdent effectivement un cœur en tonneau β . Un des résultats les plus étonnants et intéressants de ces structures a été la découverte d'un couvercle situé au-dessus de la région du site actif. Ce couvercle étant absent chez les sortases de classes A et B, nous avons émis l'hypothèse que seules les sortases impliquées dans la formation de pili possédaient un couvercle. La résolution d'autres structures de sortases de classe A, B, C ou D permettraient de confirmer cela car, à l'heure actuelle, seules 2 structures de sortases de classe A, 2 de classe B et 3 de classe C sont disponibles.

Deux « parties » peuvent être distinguées dans ce couvercle : la partie centrale, composée de deux acides aminés (un tryptophane et un aspartate) qui viennent s'ancrer directement dans le site actif où se localisent la triade catalytique histidine-cystéine-arginine, puis les bords du couvercle, de part et d'autre de la partie centrale, dont les facteurs de température indiquent une certaine flexibilité de ces zones. De cette étude, nous avons proposé que ce couvercle n'était pas simplement une région recouvrant le site actif mais avait probablement un rôle plus important. En effet, l'aspartate du couvercle interagit avec l'arginine du site actif, l'orientant vers la cystéine catalytique. De plus, l'aspartate, aidé du tryptophane, ancre le couvercle dans le site actif. D'autre part, curieusement, ce couvercle se retrouve positionné au même endroit que le peptide LPXTG dans la structure de SrtA de *S. aureus*, bloquant ainsi l'accessibilité de cette région. Nous suggérons alors que le couvercle aurait un rôle dans le maintien des résidus catalytiques dans une position optimale pour la catalyse mais également un rôle de prévention, empêchant tous substrats inappropriés d'être catalysés ; seul le bon substrat serait capable de déplacer ce couvercle, de rendre ainsi

accessible la région du site actif, et de subir la catalyse enzymatique. En effet, la surface bactérienne est le lieu de présence d'une quantité importante de protéines de virulence portant le motif LPXTG et destinées à être accrochées sur le peptidoglycane par la sortase A. Il est ainsi concevable de penser que les sortases de classe C de *S. pneumoniae* ne doivent pas interagir avec ces dernières mais avec les protéines du pilus qui sont également présentes sur la membrane bactérienne. Le couvercle pourrait alors être une « stratégie » pour éviter aux sortases impliquées dans la formation du pilus d'être détournées de leur fonction et d'interagir avec les protéines de virulence portant un motif LPXTG. La présence d'un couvercle couvrant le site actif existe également chez d'autres transpeptidases. C'est le cas chez les Penicillin Binding Proteins (PBP), enzymes intervenant dans les dernières étapes de la synthèse du peptidoglycane. Le domaine transpeptidase des PBP reconnaît un peptide au niveau de la cavité du site actif et catalyse la formation de ponts peptidiques. Une région, flexible, est notamment présente à proximité du site actif et intervient dans un phénomène de réorganisation structurale lors d'une interaction de la protéine avec un pseudo-substrat de la réaction. En effet, cette boucle flexible maintient le site actif dans une conformation fermée en absence de ligand (grâce à un résidu asparagine), mais s'ouvre en présence d'un pseudo-substrat, permettant alors la formation d'un acyl-enzyme. Les PBP et les sortases étant similaires en termes de réaction catalysée ainsi qu'en substrat reconnu, on pourrait proposer un mécanisme commun de régulation d'accès à leur site actif : en absence de substrats (chaîne peptidique pour les PBP et peptide LPXTG pour les sortases), la gorge catalytique est maintenue dans une conformation fermée à travers des interactions spécifiques générées par les résidus du couvercle. Néanmoins, lorsque le substrat devient disponible (c'est-à-dire lorsque la biosynthèse du peptidoglycane est requise pour les PBP ou lorsque la formation de pilus est requise pour les sortases), le couvercle s'ouvre, potentiellement à travers des interactions avec les substrats. Ces observations supportant l'idée d'un changement conformationnel induite par liaison au substrat devraient par conséquent être pris en compte lors de la conception d'inhibiteurs de ces enzymes.

L'étude de la région du site actif (couvercle + triade catalytique) a pu être approfondie grâce à l'étude de mutants de SrtC-1. Nous avons ainsi pu montrer que l'interaction entre l'aspartate du couvercle et l'arginine du site actif était importante pour stabiliser l'enzyme. En outre, nous avons montré que les résidus histidine et arginine du site actif sont essentiels pour que la catalyse enzymatique puisse se faire. Des études RMN ou de modélisation moléculaire seraient intéressantes à mener afin de confirmer la flexibilité du couvercle. Dans cette

perspective, une collaboration avec Emmanuel Giudice de l'Université de Rennes a été initiée.

Les sortases formant les pili : une cible thérapeutique potentielle ?

Nos études ainsi que les analyses de Lemieux et Falker (Lemieux *et al.*, 2008 ; Falker *et al.*, 2008) montrent une redondance d'activité de nos trois sortases ce qui laisse supposer que cibler les trois sortases, et non qu'une, serait plus judicieux. De plus, la présence d'un couvercle recouvrant et bloquant physiquement l'accès au site actif soulève la question d'un bon candidat pour le développement d'antibiotiques. En effet, d'après nos résultats ce couvercle serait fermé chez les sortases en absence de substrat et s'ouvrirait lors de la reconnaissance de celui-ci. Une molécule devrait ainsi mimer cette ouverture du couvercle pour pouvoir accéder au site actif et inhiber l'enzyme. Une quantité importante d'essais de cocrystallisation et de trempage avec des peptides mimant le LPXTG mais aussi avec un inhibiteur de protéase à cystéine (E64) ont été réalisés durant cette thèse afin d'obtenir une structure de sortase complexée avec un ligand. Cependant, le site actif était toujours fermé. Ces observations confirmeraient l'hypothèse d'un site actif inaccessible à un inhibiteur potentiel. Toutefois, nous avons obtenu récemment une structure complexée avec la leupeptine, inhibiteur connu de protéases à cystéine. Cette molécule, plus petite que les peptides, est donc capable de « s'engouffrer » dans le site actif, sans déplacer le couvercle. Ces résultats indiqueraient finalement qu'une petite molécule synthétique pourrait inhiber l'activité sortasique et que le développement d'antibiotiques dirigés contre les sortases serait potentiellement une piste à creuser. Il faut néanmoins noter que, même si les pili ont été montrés comme des facteurs de virulence chez le pneumocoque, ils ne se sont présents que chez 30% des souches. Cibler les sortases viendrait donc en complément de médicaments dirigés contre d'autres facteurs de virulence.

Protéine RrgB et complexe SrtC-1/RrgB, vers une structure ?

Des nombreux essais ont été menés dans l'optique de résoudre la structure de RrgB, tant en termes de criblage et optimisation de conditions de cristallisations, qu'en terme de collecte de jeux de données. En effet, les cristaux de RrgB diffractent assez mal et il est difficile de trouver un cristal diffractant à une résolution correcte. D'autre part, du fait de paramètres de maille assez grand, une collecte avec un pas d'oscillation de maximum 0,5° doit se faire, demandant ainsi un temps de collecte conséquent. Néanmoins, deux jeux (un natif et un

séléné à 3.7 Å et 3 Å respectivement) ont été enregistrés mais ne permettent encore pas de résoudre la structure. La recherche de cristaux diffractant à une meilleure résolution se poursuit. RrgB étant facilement protéolysée et instable dans le temps, il serait aussi intéressant de faire une protéolyse ménagée de RrgB afin d'isoler une partie stable de la protéine, propice à une meilleure cristallisation. Obtenir la structure de RrgB permettrait ainsi de mieux comprendre l'arrangement du pilus si tenté est que l'empilement cristallin mime la fibre.

De récentes études ont montré que la polymérisation du pilus se réalise à travers la formation d'un intermédiaire acyl-enzyme entre la piline et la sortase confortant ainsi nos travaux préliminaires sur le complexe SrtC-1/RrgB. Des études structurales sur un tel complexe seraient intéressantes à poursuivre. Obtenir sa structure tridimensionnelle permettrait effectivement, d'une part, de soulever des doutes concernant le mécanisme catalytique des sortases et, d'autre part, d'observer le comportement de ce fameux couvercle.

Vers encore une meilleure compréhension de la biogenèse des pili ?

La technique originale de formation de fibres *in vitro* que nous avons développée au laboratoire pourrait être appliquée à l'étude du pili dans son ensemble. En effet, pour le moment, aucun modèle sur l'assemblage des pili pneumococciques a été suggéré. En particulier, le mécanisme d'ancrage des pilines mineures RrgA et RrgC n'a pas été élucidé ainsi que le mécanisme d'ancrage du pilus sur le peptidoglycane. De même, l'implication des trois sortases dans l'assemblage en entier du pilus est encore mal analysée. Nous pourrions, par exemple, analyser l'ancrage des fibres de RrgB sur le lipide II afin de nous donner une idée des sortases (SrtC-1, SrtC-2, SrtC-3 ou SrtA) impliquées dans l'attachement du pilus sur le peptidoglycane. De plus, il serait intéressant de réussir à greffer RrgA et/ou RrgC sur les fibres de RrgB par le biais des sortases : d'une part cela permettrait d'analyser le ou les sortases responsables de l'attachement des protéines mineures sur le pilus et d'autre part, en fonction des mélanges réalisés, regarder si il serait possible de mettre en évidence une chronologie d'assemblage.

Les mécanismes de formation des pilus chez les bactéries à Gram-positif n'en sont qu'à leur début. De réels progrès ont d'or et déjà été réalisés, mais tout n'est pas encore expliqué. De nombreux travaux sont encore nécessaires avant l'élaboration d'un traitement efficace dirigé contre les pili. Cette course à l'armement initiée entre l'homme et la bactérie pousse la communauté scientifique à faire preuve d'ingéniosité dans la compréhension de la biogenèse de ces organelles.

Références Bibliographiques

A

AlonsoDeVelasco, E., Verheul, A.F., Verhoef, J., and Snippe, H. (1995). *Streptococcus pneumoniae*: virulence factors, pathogenesis, and vaccines. *Microbiol Rev* 59, 591-603.

Andersson, B., Dahmen, J., Frejd, T., Leffler, H., Magnusson, G., Noori, G., and Eden, C.S. (1983). Identification of an active disaccharide unit of a glycoconjugate receptor for pneumococci attaching to human pharyngeal epithelial cells. *J Exp Med* 158, 559-570.

Andersson, B., Eriksson, B., Falsen, E., Fogh, A., Hanson, L.A., Nylen, O., Peterson, H., and Svanborg Eden, C. (1981). Adhesion of *Streptococcus pneumoniae* to human pharyngeal epithelial cells in vitro: differences in adhesive capacity among strains isolated from subjects with otitis media, septicemia, or meningitis or from healthy carriers. *Infect Immun* 32, 311-317.

Andrade, M.A., Ciccarelli, F.D., Perez-Iratxeta, C., and Bork, P. (2002). NEAT: a domain duplicated in genes near the components of a putative Fe³⁺ siderophore transporter from Gram-positive pathogenic bacteria. *Genome Biol* 3, RESEARCH0047.

Andrew, P.W., Mitchell, T.J., and Morgan, P.J. (1997). Relationship of structure to function in pneumolysin. *Microb Drug Resist* 3, 11-17.

Austrian, R. (1981). Some observations on the pneumococcus and on the current status of pneumococcal disease and its prevention. *Rev Infect Dis* 3 *Suppl*, S1-17.

B

Bagnoli, F., Moschioni, M., Donati, C., Dimitrovska, V., Ferlenghi, I., Facciotti, C., Muzzi, A., Giusti, F., Emolo, C., Sinisi, A., *et al.* (2008). A second pilus type in *Streptococcus pneumoniae* is prevalent in emerging serotypes and mediates adhesion to host cells. *J Bacteriol* 190, 5480-5492.

Balachandran, P., Brooks-Walter, A., Virolainen-Julkunen, A., Hollingshead, S.K., and Briles, D.E. (2002). Role of pneumococcal surface protein C in nasopharyngeal carriage and pneumonia and its ability to elicit protection against carriage of *Streptococcus pneumoniae*. *Infect Immun* 70, 2526-2534.

Barnett, T.C., and Scott, J.R. (2002). Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. *J Bacteriol* 184, 2181-2191.

Barocchi, M.A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernebro, J., Moschioni, M., Masignani, V., *et al.* (2006). A pneumococcal pilus influences virulence and host inflammatory responses. *Proc Natl Acad Sci U S A* 103, 2857-2862.

Basset, A., Trzcinski, K., Hermos, C., O'Brien, K.L., Reid, R., Santosham, M., McAdam, A.J., Lipsitch, M., and Malley, R. (2007). Association of the pneumococcal pilus with certain capsular serotypes but not with increased virulence. *J Clin Microbiol* 45, 1684-1689.

Bentley, M.L., Lamb, E.C., and McCafferty, D.G. (2008). Mutagenesis studies of substrate recognition and catalysis in the sortase A transpeptidase from *Staphylococcus aureus*. *J Biol Chem* 283, 14762-14771.

Berry, A.M., Lock, R.A., Hansman, D., and Paton, J.C. (1989). Contribution of autolysin to virulence of *Streptococcus pneumoniae*. *Infect Immun* 57, 2324-2330.

Berry, A.M., and Paton, J.C. (2000). Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins. *Infect Immun* 68, 133-140.

Bierne, H., Garandeau, C., Pucciarelli, M.G., Sabet, C., Newton, S., Garcia-del Portillo, F., Cossart, P., and Charbit, A. (2004). Sortase B, a new class of sortase in *Listeria monocytogenes*. *J Bacteriol* 186, 1972-1982.

Bierne, H., Mazmanian, S.K., Trost, M., Pucciarelli, M.G., Liu, G., Dehoux, P., Jansch, L., Garcia-del Portillo, F., Schneewind, O., Cossart, P., and European *Listeria* Genome, C. (2002). Inactivation of the *srtA* gene in *Listeria monocytogenes* inhibits anchoring of surface proteins and affects virulence. *Mol Microbiol* 43, 869-881.

Bolken, T.C., Franke, C.A., Jones, K.F., Zeller, G.O., Jones, C.H., Dutton, E.K., and Hruby, D.E. (2001). Inactivation of the *srtA* gene in *Streptococcus gordonii* inhibits cell wall anchoring of surface proteins and decreases in vitro and in vivo adhesion. *Infect Immun* 69, 75-80.

Boulnois, G.J. (1992). Pneumococcal proteins and the pathogenesis of disease caused by *Streptococcus pneumoniae*. *J Gen Microbiol* 138, 249-259.

Bricker, A.L., and Camilli, A. (1999). Transformation of a type 4 encapsulated strain of *Streptococcus pneumoniae*. *FEMS Microbiol Lett* 172, 131-135.

Bricogne, G., Vonnrhein, C., Flensburg, C., Schiltz, M., and Paciorek, W. (2003). Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr* 59, 2023-2030.

Briles, D.E., Hollingshead, S.K., Paton, J.C., Ades, E.W., Novak, L., van Ginkel, F.W., and Benjamin, W.H., Jr. (2003). Immunizations with pneumococcal surface protein A and pneumolysin are protective against pneumonia in a murine model of pulmonary infection with *Streptococcus pneumoniae*. *J Infect Dis* 188, 339-348.

Briles, D.E., Yother, J., and McDaniel, L.S. (1988). Role of pneumococcal surface protein A in the virulence of *Streptococcus pneumoniae*. *Rev Infect Dis* 10 Suppl 2, S372-374.

C

Casal, J. (1982). Antimicrobial susceptibility of *Streptococcus pneumoniae*: serotype distribution of penicillin-resistant strains in Spain. *Antimicrob Agents Chemother* 22, 222-225.

Centers for Disease Control and Prevention. (2005). Direct and indirect effects of routine vaccination of children with 7-valent pneumococcal conjugate vaccine on incidence of invasive pneumococcal disease--United States, 1998-2003. *MMWR Morb Mortal Wkly Rep* 54, 893-897.

Cisar, J.O., Vatter, A.E., Clark, W.B., Curl, S.H., Hurst-Calderone, S., and Sandberg, A.L. (1988). Mutants of *Actinomyces viscosus* T14V lacking type 1, type 2, or both types of fimbriae. *Infect Immun* 56, 2984-2989.

Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-763.

Cossart, P., and Jonquieres, R. (2000). Sortase, a universal target for therapeutic agents against gram-positive bacteria? *Proc Natl Acad Sci U S A* 97, 5013-5015.

D

Deibel, R., and Seeley, H. (1974). Streptococcaceae fam. nov. *Bergey's Manual of Determinative Bacteriology*, 8, 490-509.

Douglas, R.M., Paton, J.C., Duncan, S.J., and Hansman, D.J. (1983). Antibody response to pneumococcal vaccination in children younger than five years of age. *J Infect Dis* 148, 131-137.

Dramsi, S., Trieu-Cuot, P., and Bierne, H. (2005). Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. *Res Microbiol* 156, 289-297.

Driessen, A. J., and Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem* 77, 643-67.

E

Eberhardt, C., Kuerschner, L., and Weiss, D.S. (2003). Probing the catalytic activity of a cell division-specific transpeptidase in vivo with beta-lactams. *J Bacteriol* 185, 3726-3734.

Edwards, A.M., Manetti, A.G., Falugi, F., Zingaretti, C., Capo, S., Buccato, S., Bensi, G., Telford, J.L., Margarit, I., and Grandi, G. (2008). Scavenger receptor gp340 aggregates group A streptococci by binding pili. *Mol Microbiol* 68, 1378-1394.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.

F

Falker, S., Nelson, A.L., Morfeldt, E., Jonas, K., Hultenby, K., Ries, J., Melefors, O., Normark, S., and Henriques-Normark, B. (2008). Sortase-mediated assembly and surface topology of adhesive pneumococcal pili. *Mol Microbiol* 70, 595-607.

Fives-Taylor, P.M., and Thompson, D.W. (1985). Surface properties of *Streptococcus sanguis* FW213 mutants nonadherent to saliva-coated hydroxyapatite. *Infect Immun* 47, 752-759.

Frankel, B.A., Kruger, R.G., Robinson, D.E., Kelleher, N.L., and McCafferty, D.G. (2005). *Staphylococcus aureus* sortase transpeptidase SrtA: insight into the kinetic mechanism and evidence for a reverse protonation catalytic mechanism. *Biochemistry* 44, 11188-11200.

Frankel, B.A., Tong, Y., Bentley, M.L., Fitzgerald, M.C., and McCafferty, D.G. (2007). Mutational analysis of active site residues in the *Staphylococcus aureus* transpeptidase SrtA. *Biochemistry* 46, 7269-7278.

G

Gaspar, A.H., and Ton-That, H. (2006). Assembly of distinct pilus structures on the surface of *Corynebacterium diphtheriae*. *J Bacteriol* 188, 1526-1533.

Geelen, S., Bhattacharyya, C., and Tuomanen, E. (1993). The cell wall mediates pneumococcal attachment to and cytopathology in human endothelial cells. *Infect Immun* 61, 1538-1543.

Geerlof, A., Brown, J., Coutard, B., Egloff, M.P., Enguita, F.J., Fogg, M.J., Gilbert, R.J., Groves, M.R., Haouz, A., Nettleship, J.E., Nordlund, P., Owens, R.J., Ruff, M., Sainsbury, S., Svergun, D.I., and Wilmanns, M. (2006). The impact of protein characterization in structural proteomics. *Acta Crystallogr D Biol Crystallogr* 62, 1125-1136

Geslin, P., Buu-Hoi, A., Fremaux, A., and Acar, J.F. (1992). Antimicrobial resistance in *Streptococcus pneumoniae*: an epidemiological survey in France, 1970-1990. *Clin Infect Dis* 15, 95-98.

Gianfaldoni, C., Censini, S., Hilleringmann, M., Moschioni, M., Facciotti, C., Pansegrau, W., Masignani, V., Covacci, A., Rappuoli, R., Barocchi, M.A., and Ruggiero, P. (2007). *Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge. *Infect Immun* 75, 1059-1062.

H

Hammerschmidt, S., Bethe, G., Remane, P.H., and Chhatwal, G.S. (1999). Identification of pneumococcal surface protein A as a lactoferrin-binding protein of *Streptococcus pneumoniae*. *Infect Immun* 67, 1683-1687.

Hava, D.L., and Camilli, A. (2002). Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol* 45, 1389-1406.

Henrichsen, J. (1995). Six newly recognized types of *Streptococcus pneumoniae*. *J Clin Microbiol* 33, 2759-2762.

Hicks, L.A., Harrison, L.H., Flannery, B., Hadler, J.L., Schaffner, W., Craig, A.S., Jackson, D., Thomas, A., Beall, B., Lynfield, R., *et al.* (2007). Incidence of pneumococcal disease due

to non-pneumococcal conjugate vaccine (PCV7) serotypes in the United States during the era of widespread PCV7 vaccination, 1998-2004. *J Infect Dis* 196, 1346-1354.

Hilleringmann, M., Giusti, F., Baudner, B.C., Maignani, V., Covacci, A., Rappuoli, R., Barocchi, M.A., and Ferlenghi, I. (2008). Pneumococcal pili are composed of protofilaments exposing adhesive clusters of Rrg A. *PLoS Pathog* 4, e1000026.

J

Jacobs, M.R., Koornhof, H.J., Robins-Browne, R.M., Stevenson, C.M., Vermaak, Z.A., Freiman, I., Miller, G.B., Witcomb, M.A., Isaacson, M., Ward, J.I., and Austrian, R. (1978). Emergence of multiply resistant pneumococci. *N Engl J Med* 299, 735-740.

Jonsson, I.M., Mazmanian, S.K., Schneewind, O., Bremell, T., and Tarkowski, A. (2003). The role of *Staphylococcus aureus* sortase A and sortase B in murine arthritis. *Microbes Infect* 5, 775-780.

K

Kabsch (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. pp. 795-800.

Kang, H.J., Coulibaly, F., Clow, F., Proft, T., and Baker, E.N. (2007). Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. *Science* 318, 1625-1628.

Kelly, T., Dillard, J.P., and Yother, J. (1994). Effect of genetic switching of capsular type on virulence of *Streptococcus pneumoniae*. *Infect Immun* 62, 1813-1819.

Kharat, A.S., and Tomasz, A. (2003). Inactivation of the *srtA* gene affects localization of surface proteins and decreases adhesion of *Streptococcus pneumoniae* to human pharyngeal cells in vitro. *Infect Immun* 71, 2758-2765.

Koebnik, R. (2001). The role of bacterial pili in protein and DNA translocation. *Trends Microbiol* 9, 586-590.

Krishnan, V., Gaspar, A.H., Ye, N., Mandlik, A., Ton-That, H., and Narayana, S.V. (2007). An IgG-like domain in the minor pilin GBS52 of *Streptococcus agalactiae* mediates lung epithelial cell adhesion. *Structure* 15, 893-903.

L

Lalioui, L., Pellegrini, E., Dramsi, S., Baptista, M., Bourgeois, N., Doucet-Populaire, F., Rusniok, C., Zouine, M., Glaser, P., Kunst, F., *et al.* (2005). The SrtA Sortase of *Streptococcus agalactiae* is required for cell wall anchoring of proteins containing the LPXTG motif, for adhesion to epithelial cells, and for colonization of the mouse intestine. *Infect Immun* 73, 3342-3350.

Lauer, P., Rinaudo, C.D., Soriani, M., Margarit, I., Maione, D., Rosini, R., Taddei, A.R., Mora, M., Rappuoli, R., Grandi, G., and Telford, J.L. (2005). Genome analysis reveals pili in Group B Streptococcus. *Science* 309, 105.

LeMieux, J., Woody, S., and Camilli, A. (2008). Roles of the sortases of *Streptococcus pneumoniae* in assembly of the RlrA pilus. *J Bacteriol* 190, 6002-6013.

Lepoutre, A., Varon, E., Georges, S., Gutmann, L., and Levy-Bruhl, D. (2008). Impact of infant pneumococcal vaccination on invasive pneumococcal diseases in France, 2001-2006. *Euro Surveill* 13.

Levesque, C., Vadeboncoeur, C., Chandad, F., and Frenette, M. (2001). *Streptococcus salivarius* fimbriae are composed of a glycoprotein containing a repeated motif assembled into a filamentous nondissociable structure. *J Bacteriol* 183, 2724-2732.

Levy, D. (1998). Epidémiologie de la flore nasopharyngée au cours des otites moyennes aiguës de l'enfant en 1997. (*Med. Mal. Infect.*), pp. 8-22.

Lu, C., Zhu, J., Wang, Y., Umeda, A., Cowmeadow, R.B., Lai, E., Moreno, G.N., Person, M.D., and Zhang, Z. (2007). *Staphylococcus aureus* sortase A exists as a dimeric protein in vitro. *Biochemistry* 46, 9346-9354.

M

MacLeod, C.M. (1945). Prevention of pneumococcal pneumoniae by immunization with specific capsular polysaccharides. *J. Exp. Med.* 82, 445-465.

Maisey, H.C., Hensler, M., Nizet, V., and Doran, K.S. (2007). Group B streptococcal pilus proteins contribute to adherence to and invasion of brain microvascular endothelial cells. *J Bacteriol* 189, 1464-1467.

Mandlik, A., Das, A., and Ton-That, H. (2008). The molecular switch that activates the cell wall anchoring step of pilus assembly in gram-positive bacteria. *Proc Natl Acad Sci U S A* 105, 14147-14152.

Mandlik, A., Swierczynski, A., Das, A., and Ton-That, H. (2007). *Corynebacterium diphtheriae* employs specific minor pilins to target human pharyngeal epithelial cells. *Mol Microbiol* 64, 111-124.

Manetti, A.G., Zingaretti, C., Falugi, F., Capo, S., Bombaci, M., Bagnoli, F., Gambellini, G., Bensi, G., Mora, M., Edwards, A.M., *et al.* (2007). *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. *Mol Microbiol* 64, 968-983.

Manzano, C., Contreras-Martel, C., El Mortaji, L., Izoré, T., Fenel, D., Vernet, T., Schoehn, G., Di Guilmi, A.M., and Dessen, A. (2008). Sortase-mediated pilus fiber biogenesis in *Streptococcus pneumoniae*. *Structure* 16, 1838-1848.

Maresso, A.W., Chapa, T.J., and Schneewind, O. (2006). Surface protein IsdC and Sortase B are required for heme-iron scavenging of *Bacillus anthracis*. *J Bacteriol* 188, 8145-8152.

- Maresso, A.W., and Schneewind, O. (2006). Iron acquisition and transport in *Staphylococcus aureus*. *Biometals* 19, 193-203.
- Marraffini, L.A., and Schneewind, O. (2006). Targeting proteins to the cell wall of sporulating *Bacillus anthracis*. *Mol Microbiol* 62, 1402-1417.
- Marraffini, L.A., Ton-That, H., Zong, Y., Narayana, S.V., and Schneewind, O. (2004). Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. A conserved arginine residue is required for efficient catalysis of sortase A. *J Biol Chem* 279, 37763-37770.
- Marton, A., Gulyas, M., Munoz, R., and Tomasz, A. (1991). Extremely high incidence of antibiotic resistance in clinical isolates of *Streptococcus pneumoniae* in Hungary. *J Infect Dis* 163, 542-548.
- Mattick, J.S. (2002). Type IV pili and twitching motility. *Annu Rev Microbiol* 56, 289-314.
- Mazmanian, S.K., Skaar, E.P., Gaspar, A.H., Humayun, M., Gornicki, P., Jelenska, J., Joachmiak, A., Missiakas, D.M., and Schneewind, O. (2003). Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* 299, 906-909.
- Mazmanian, S.K., Ton-That, H., Su, K., and Schneewind, O. (2002). An iron-regulated sortase anchors a class of surface protein during *Staphylococcus aureus* pathogenesis. *Proc Natl Acad Sci U S A* 99, 2293-2298.
- McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 61, 458-464.
- McNab, R., Forbes, H., Handley, P.S., Loach, D.M., Tannock, G.W., and Jenkinson, H.F. (1999). Cell wall-anchored CshA polypeptide (259 kilodaltons) in *Streptococcus gordonii* forms surface fibrils that confer hydrophobic and adhesive properties. *J Bacteriol* 181, 3087-3095.
- McPherson, A., Malkin, A.J., and Kuznetsov, Y.G. (1995). The science of macromolecular crystallization. *Structure* 3, 759-768.
- Mosser, J.L., and Tomasz, A. (1970). Choline-containing teichoic acid as a structural component of pneumococcal cell wall and its role in sensitivity to lysis by an autolytic enzyme. *J Biol Chem* 245, 287-98.
- Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A.G., Maggi, T., Taddei, A.R., Grandi, G., and Telford, J.L. (2005). Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. *Proc Natl Acad Sci U S A* 102, 15641-15646.
- Morris, R.J., Perrakis, A., and Lamzin, V.S. (2003). ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol* 374, 229-244.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53, 240-255.

Musher, D.M. (1992). Infections caused by *Streptococcus pneumoniae*: clinical spectrum, pathogenesis, immunity, and treatment. *Clin Infect Dis* 14, 801-807.

N

Naik, M.T., Suree, N., Ilangoan, U., Liew, C.K., Thieu, W., Campbell, D.O., Clemens, J.J., Jung, M.E., and Clubb, R.T. (2006). *Staphylococcus aureus* Sortase A transpeptidase. Calcium promotes sorting signal binding by altering the mobility and structure of an active site loop. *J Biol Chem* 281, 1817-1826.

Navarre, W.W., and Schneewind, O. (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* 63, 174-229.

Nelson, A.L., Ries, J., Bagnoli, F., Dahlberg, S., Falker, S., Rounioja, S., Tschop, J., Morfeldt, E., Ferlenghi, I., Hilleringmann, M., Holden, D. W., Rappuoli, R., Normark, S., Barocchi, M. A., Henriques-Normark, B. (2007). RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol Microbiol* 66, 329-340.

O

Ogunniyi, A.D., Grabowicz, M., Briles, D.E., Cook, J., and Paton, J.C. (2007). Development of a vaccine against invasive pneumococcal disease based on combinations of virulence proteins of *Streptococcus pneumoniae*. *Infect Immun* 75, 350-357.

Osaki, M., Takamatsu, D., Shimoji, Y., and Sekizaki, T. (2002). Characterization of *Streptococcus suis* genes encoding proteins homologous to sortase of gram-positive bacteria. *J Bacteriol* 184, 971-982.

P

Pallen, M.J., Lam, A.C., Antonio, M., and Dunbar, K. (2001). An embarrassment of sortases - a richness of substrates? *Trends Microbiol* 9, 97-102.

Pantoliano, M.W., Petrella, E.C., Kwasnoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B.A., Lane, P., and Salemme, F.R. (2001). High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* 6, 429-440

Pasteur, L., Note sur la maladie nouvelle provoquée par la salive d'un enfant mort de la rage. *Bulletin de l'Académie de Médecine*. 10, 94-103.

Park, I.H., Pritchard, D.G., Cartee, R., Brandao, A., Brandileone, M.C., and Nahm, M.H. (2007). Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* 45, 1225-1233.

Paton, J.C., and Ferrante, A. (1983). Inhibition of human polymorphonuclear leukocyte respiratory burst, bactericidal activity, and migration by pneumolysin. *Infect Immun* 41, 1212-1216.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6, 458-463.

R

Rayner, C.F., Jackson, A.D., Rutman, A., Dewar, A., Mitchell, T.J., Andrew, P.W., Cole, P.J., and Wilson, R. (1995). Interaction of pneumolysin-sufficient and -deficient isogenic variants of *Streptococcus pneumoniae* with human respiratory mucosa. *Infect Immun* 63, 442-447.

Robbins, J.B., Austrian, R., Lee, C.J., Rastogi, S.C., Schiffman, G., Henrichsen, J., Makela, P.H., Broome, C.V., Facklam, R.R., Tiesjema, R.H., and Parke J.C. (1983). Considerations for formulating the second-generation pneumococcal capsular polysaccharide vaccine with emphasis on the cross-reactive types within groups. *J Infect Dis* 148, 1136-1159.

Rosch, J.W., Mann, B., Thornton, J., Sublett, J., and Tuomanen, E. (2008). Convergence of regulatory networks on the pilus locus of *Streptococcus pneumoniae*. *Infect Immun* 76, 3187-3196.

Rosini, R., Rinaudo, C.D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzi, C., Brettoni, C., *et al.* (2006). Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. *Mol Microbiol* 61, 126-141.

Rubins, J.B., Duane, P.G., Charboneau, D., and Janoff, E.N. (1992). Toxicity of pneumolysin to pulmonary endothelial cells in vitro. *Infect Immun* 60, 1740-1746.

S

Sauer, F.G., Remaut, H., Hultgren, S.J., and Waksman, G. (2004). Fiber assembly by the chaperone-usher pathway. *Biochim Biophys Acta* 1694, 259-267.

Schneewind, O., Mihaylova-Petkov, D., Model, P. (1993). Cell wall sorting signals in surface proteins of gram-positive bacteria. *EMBO J* 12, 4803-4811.

Schneewind, O., Model, P., Fischetti, V.A. (1992). Sorting of protein A to the staphylococcal cell wall. *Cell* 70, 267-281.

Schneewind, O., Fowler, A., and Faull, K.F. (1995). Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. *Science* 268, 103-106.

Scott, J.R., and Zahner, D. (2006). Pili with strong attachments: Gram-positive bacteria do it differently. *Mol Microbiol* 62, 320-330.

Soto, G.E., and Hultgren, S.J. (1999). Bacterial adhesins: common themes and variations in architecture and assembly. *J Bacteriol* 181, 1059-1071.

Steinfert, C., Wilson, R., Mitchell, T., Feldman, C., Rutman, A., Todd, H., Sykes, D., Walker, J., Saunders, K., Andrew, P.W., and *et al.* (1989). Effect of *Streptococcus pneumoniae* on human respiratory epithelium in vitro. *Infect Immun* 57, 2006-2013.

Stenberg, G.M., (1881). A fatal form of septicæmia in the rabbit, produced by the subcutaneous injection of human saliva. (Annual Reports of the National Board of Health), pp. 87-108.

Storoni, L.C., McCoy, A.J., and Read, R.J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* 60, 432-438.

Swaminathan, A., Mandlik, A., Swierczynski, A., Gaspar, A., Das, A., and Ton-That, H. (2007). Housekeeping sortase facilitates the cell wall anchoring of pilus polymers in *Corynebacterium diphtheriae*. *Mol Microbiol* 66, 961-974.

T

Telford, J.L., Barocchi, M.A., Margarit, I., Rappuoli, R., Grandi, G. (2006). Pili in gram-positive pathogens. *Nat Rev Microbiol* 4, 509-519

Ton-That, H., Faull, K.F., and Schneewind, O. (1997). Anchor structure of staphylococcal surface proteins. A branched peptide that links the carboxyl terminus of proteins to the cell wall. *J Biol Chem* 272, 22285-22292.

Ton-That, H., Liu, G., Mazmanian, S.K., Faull, K.F., and Schneewind, O. (1999). Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc Natl Acad Sci U S A* 96, 12424-12429.

Ton-That, H., Mazmanian, S.K., Faull, K.F., and Schneewind, O. (2000). Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates. *J Biol Chem* 275, 9876-9881.

Ton-That, H., and Schneewind, O. (2003). Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol Microbiol* 50, 1429-1438.

Ton-That, H., Marraffini, L.A., and Schneewind, O. (2004). Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol Microbiol* 53, 251-261.

Ton-That, H., Mazmanian, S.K., Alksne, L., and Schneewind, O. (2002). Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Cysteine 184 and histidine 120 of sortase form a thiolate-imidazolium ion pair for catalysis. *J Biol Chem* 277, 7447-7452.

Ton-That, H., and Schneewind, O. (2004). Assembly of pili in Gram-positive bacteria. *Trends Microbiol* 12, 228-234.

V

Varon, E. (2001). Severe pneumococcal infections: virulence aspects. *Arch Pediatr* 8 Suppl 4, 752s-756s.

Vernet, T., Tessier, D.C., Chatellier, J., Plouffe, C., Lee, T.S., Thomas, D.Y., Storer, A.C., and Menard, R. (1995). Structural and functional roles of asparagine 175 in the cysteine protease papain. *J Biol Chem* 270, 16645-16652.

W

Watson, D.A., and Musher, D.M. (1990). Interruption of capsule production in *Streptococcus pneumoniae* serotype 3 by insertion of transposon Tn916. *Infect Immun* 58, 3135-3138.

Weiss, W.J., Lenoy, E., Murphy, T., Tardio, L., Burgio, P., Projan, S.J., Schneewind, O., and Alksne, L. (2004). Effect of *srtA* and *srtB* gene expression on the virulence of *Staphylococcus aureus* in animal models of infection. *J Antimicrob Chemother* 53, 480-486.

Willcox, M.D., and Drucker, D.B. (1989). Surface structures, co-aggregation and adherence phenomena of *Streptococcus oralis* and related species. *Microbios* 59, 19-29.

Willcox, M.D., Wyatt, J.E., and Handley, P.S. (1989). A comparison of the adhesive properties and surface ultrastructure of the fibrillar *Streptococcus sanguis* 12 and an adhesion deficient non-fibrillar mutant 12 na. *J Appl Bacteriol* 66, 291-299.

Winter, A.J., Comis, S.D., Osborne, M.P., Tarlow, M.J., Stephen, J., Andrew, P.W., Hill, J., Mitchell, T.J. (1997). A role for pneumolysin but not neuraminidase in the hearing loss and cochlear damage induced by experimental pneumococcal meningitis in guinea pigs. *Infect Immun* 65, 4411-4418

Y

Yanagawa, R., and Honda, E. (1976). Presence of pili in species of human and animal parasites and pathogens of the genus *Corynebacterium*. *Infect Immun* 13, 1293-1295.

Yanagawa, R., Otsuki, K., and Tokui, T. (1968). Electron microscopy of fine structure of *Corynebacterium renale* with special reference to pili. *Jpn J Vet Res* 16, 31-37.

Yeung, M.K., Donkersloot, J.A., Cisar, J.O., and Ragsdale, P.A. (1998). Identification of a gene involved in assembly of *Actinomyces naeslundii* T14V type 2 fimbriae. *Infect Immun* 66, 1482-1491.

Yeung, M.K., and Ragsdale, P.A. (1997). Synthesis and function of *Actinomyces naeslundii* T14V type 1 fimbriae require the expression of additional fimbria-associated genes. *Infect Immun* 65, 2629-2639.

Z

Zahner, D., and Scott, J.R. (2008). SipA is required for pilus formation in *Streptococcus pyogenes* serotype M3. *J Bacteriol* 190, 527-535.

Zhang, J.R., Mostov, K.E., Lamm, M.E., Nanno, M., Shimida, S., Ohwaki, M., and Tuomanen, E. (2000). The polymeric immunoglobulin receptor translocates pneumococci across human nasopharyngeal epithelial cells. *Cell* 102, 827-837.

Zhang, R., Wu, R., Joachimiak, G., Mazmanian, S.K., Missiakas, D.M., Gornicki, P., Schneewind, O., and Joachimiak, A. (2004). Structures of sortase B from *Staphylococcus aureus* and *Bacillus anthracis* reveal catalytic amino acid triad in the active site. *Structure* 12, 1147-1156.

Zong, Y., Bice, T.W., Ton-That, H., Schneewind, O., and Narayana, S.V. (2004a). Crystal structures of *Staphylococcus aureus* sortase A and its substrate complex. *J Biol Chem* 279, 31383-31389.

Zong, Y., Mazmanian, S.K., Schneewind, O., and Narayana, S.V. (2004b). The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the *Staphylococcus aureus* cell wall. *Structure* 12, 105-112.

Annexes

Annexe A

Mutants de SrtC-1	
Mutations	Séquences des amorces 5'→ 3'
C193A	GACCTTGCTGACT <u>GCT</u> ACGCCATACATG CATGTATGGCGT <u>AGC</u> AGTCAGCAAGGTC
H131D	GGTGATTACGGCAG <u>ATA</u> CAGGTTTGCC GGCAAACCTGT <u>ATCT</u> GCCGTATCACC
R202E	GATCAATACCCATG <u>AGC</u> TATTGGTTCGGGG CCCCGAACCAATAG <u>CTC</u> ATGGGTATTGATC
D58GW60G	AATGTAGTGAGTGGC <u>GTCCTGGGTC</u> GGAAGAAATGAAG CTTCATTTCTTCGACCCAGG <u>ACCGCC</u> ACTCACTACATT

Tableau 5 : Amorces nucléotidiques utilisées pour réaliser les mutants de SrtC-1. Les bases soulignées sont celles qui sont mutées.

<i>SrtC-1Cys193Ala</i>		<i>SrtC-1His131Asp et SrtC-1Arg202Glu</i>		<i>SrtC-1Asp58GlyTrp60Gly</i>	
95 °C	1 min	95 °C	1 min	95 °C	30 sec
95 °C	1 min	95 °C	1 min	95 °C	30 sec
55 °C	1 min	55 °C	30 sec	55 °C	1 min
68 °C	10 min	68 °C	6 min	68 °C	8 min
16 cycles		16 cycles		17 cycles	

Figure 52 : Cycles de PCR effectués pour construire les mutants de SrtC-1 et RrgB.

Mutant SrtC-1 avec le couvercle de SrtC-2	
Mutations	Séquences des amorces 5'→ 3'
SrtC-1_lid2	GGCTCAAGCCTTCAATGCGACCTTG AATAATGTAGTGAGTGGCGATCCTTG GAATTTCCACATAGCCAATCCGCTCATG GATTTCTAACATACGTGCATACTCTGC

Mutant SrtC-2 avec le couvercle de SrtC-1	
Mutations	Séquences des amorces 5'→ 3'
SrtC-2_lid1	GGCACAAGCCTTCAATGACTCTTTG AAACCATCTGAAATTCTTGATCCTTTTAC GATTTCCACATGCCCCATCCGCTCATG GACCTTTAGCATATTGGCATATTCTG

Tableau 6 : Amorces nucléotidiques utilisées pour construire les mutants SrtC-1_lid2 et SrtC-2_lid1. Ce qui est écrit et surligné en rose correspond à SrtC-2 alors que ce qui est écrit et surligné en cyan correspond à SrtC-1

A

<i>SrtC-1_lid2 et SrtC-2_lid1</i>	
95 °C	1 min
95 °C	30 sec
55 °C	30 sec
72 °C	20 sec
3 cycles	
95 °C	30 sec
60 °C	30 sec
72 °C	20 sec
27 cycles	
72 °C	10 min

B

<i>SrtC-1_lid2 et SrtC-2_lid1</i>	
95 °C	1 min
95 °C	1 min
55 °C	1 min
68 °C	22 min
22 cycles	
68 °C	18 min

Figure 53 : Cycles de PCR effectués pour construire les mutant SrtC-1_lid2 et SrtC-2_lid1. A/ Première PCR réalisée pour générer la région génomique correspondant aux couvercles de SrtC-1 et SrtC-2. B/ Seconde PCR effectuée pour construire les deux mutants.

Annexe B

Milieu de culture minimum M9 (1L)

Milieu minimum

750 mL H₂O stérile

200 mL tampon sels 5x

2 mL MgSO₄ 1 M

10 mL glucose 20 %

0.1 mL CaCl₂ 1 M

2 mL thiamine 10 %

qsp 1 L H₂O

Tampon sels 5x (500 mL)

15 g Na₂HPO₄ anhydre

7.5 g KH₂PO₄

2.5 g NH₄Cl

1.25 g NaCl

qsp 500 mL H₂O

Publications

Les travaux décrits dans ce manuscrit ont fait l'objet de deux publications :

Sortase-mediated pilus fiber biogenesis in *Streptococcus pneumoniae*. Manzano C., Contreras-Martel C., El Mortaji L., Izoré T., Fenel D., Vernet T., Schoehn G., Di Guilmi A.M., Dessen A. (2008) *Structure* **16**(12):1838-48.

Sortase activity is controlled by a flexible lid in the pilus biogenesis mechanism of Gram-positive pathogens. Manzano C., Di Guilmi A.M., Dessen A. (Soumis).

Un article additionnel a été également accepté dans J. Mol. Biol. avec la collaboration de Birgitta Henriques-Normark (Karolinska Institutet, Sweden) :

The crystal structure of *Streptococcus pneumoniae* Sortase C provides novel insights into catalysis as well as pilin substrate specificity. Neiers F., Madhurantakam C., Fälker S., Manzano C., Dessen A., Normark S., Henriques-Normark B., Achour A.

Un autre article, référant au travail de Thierry Izoré, Master II que j'ai encadré et concernant la protéine RrgA, a été soumis.

Sortase-Mediated Pilus Fiber Biogenesis in *Streptococcus pneumoniae*

Clothilde Manzano,¹ Carlos Contreras-Martel,¹ Lamy El Mortaji,² Thierry Izoré,¹ Daphna Fenel,³ Thierry Vernet,² Guy Schoehn,⁴ Anne Marie Di Guilmi,^{2,*} and Andréa Dessen^{1,*}

¹Laboratoire des Protéines Membranaires

²Laboratoire d'Ingénierie des Macromolécules

³Laboratoire de Microscopie Electronique et Structurale

Institut de Biologie Structurale Jean-Pierre Ebel, UMR 5075 (CEA, CNRS, UJF, PSB), 41 rue Jules Horowitz, F-38027 Grenoble, France

⁴Unit for Virus Host Cell Interactions, UMR5233 UJF-EMBL-CNRS, CIBB, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France

*Correspondence: diguilmi@ibs.fr (A.M.D.G.), dessen@ibs.fr (A.D.)

DOI 10.1016/j.str.2008.10.007

SUMMARY

Streptococcus pneumoniae is a pilated pathogen whose ability to circumvent vaccination and antibiotic treatment strategies is a cause of mortality worldwide. Pili play important roles in pneumococcal infection, but little is known about their biogenesis mechanism or the relationship between components of the pilus-forming machinery, which includes the fiber pilin (RrgB), two minor pilins (RrgA, RrgC), and three sortases (SrtC-1, SrtC-2, SrtC-3). Here we show that SrtC-1 is the main pilus-polymerizing transpeptidase, and electron microscopy analyses of RrgB fibers reconstituted in vitro reveal that they structurally mimic the pneumococcal pilus backbone. Crystal structures of both SrtC-1 and SrtC-3 reveal active sites whose access is controlled by flexible lids, unlike in non-pilus sortases, and suggest that substrate specificity is dictated by surface recognition coupled to lid opening. The distinct structural features of pilus-forming sortases suggest a common pilus biogenesis mechanism that could be exploited for the development of broad-spectrum antibacterials.

INTRODUCTION

Streptococcus pneumoniae, the causative agent of pneumonia, bacteremia, otitis, and meningitis, is a major cause of community-acquired illnesses, especially in the very young and the elderly, and causes over 1.6 million deaths worldwide each year (Levine et al., 2006). The spread of antibiotic-resistant strains (Bronzwaer et al., 2002), coupled to the limited efficiency and/or the restricted strain coverage of currently commercialized vaccines (Bogaert et al., 2004; Spratt and Greenwood, 2000), underline the need for the study of novel therapeutic targets for pneumococcal infection.

Secreted virulence factors and extracellular appendages are not only key factors in the initiation of bacterial infection, but also noteworthy vaccine antigens (Gianfaldoni et al., 2007; Maione et al., 2005; Mora et al., 2005). Pili, hair-like, elongated fibers associated to the bacterial cell wall, play important roles

in adhesion, biofilm formation, competence for DNA transformation, and motility, and are thus key elements of a bacterium's pathogenicity arsenal (Burrows, 2005; Manetti et al., 2007; Matlick, 2002). Pili expressed by Gram-negative pathogens have been the targets of multiple studies that have revealed that these filaments are composed of non-covalently associated subunits; coordination of subunit association involves, in many cases, strand exchange mechanisms orchestrated by periplasmic chaperones (Sauer et al., 2002; Vetsch et al., 2004). The pilus formation process in Gram-positive bacteria, however, is still poorly understood. It is widely accepted that in pathogens such as *Corynebacterium diphtheriae*, the causative agent of diphtheria, *Streptococcus pyogenes*, which causes toxic shock syndrome, and *S. pneumoniae*, pili are formed through the covalent association of subunits by sortases that are encoded within pilus-specific pathogenicity islands (Budzik et al., 2008; LeMieux et al., 2006; Mandlik et al., 2008; Mora et al., 2005; Ton-That et al., 2004; Ton-That and Schneewind, 2003). Notably, sortases are membrane-associated transpeptidases that play roles in functions ranging from iron acquisition to sporulation (Marraffini and Schneewind, 2006; Mazmanian et al., 2003). The quintessential, best-studied sortase, SrtA, catalyzes the covalent attachment of surface-exposed proteins, such as virulence factors, to the *Staphylococcus aureus* peptidoglycan (Schneewind et al., 1992). Its catalytic mechanism involves recognition of an LPXTG-like motif on the target protein followed by nucleophilic attack of the Thr-Gly covalent bond by a catalytic Cys residue. The ensuing acyl-enzyme intermediate can only be resolved by attack from a free ϵ -amino group from the secondary substrate that, in the case of SrtA, is the bridging residue of the stem peptide of the peptidoglycan (Marraffini et al., 2006; Ton-That et al., 2002). Pilus formation, however, requires dedicated sortases, and the precise roles these molecules play in the pilus biogenesis process are still poorly understood. It is of interest that the recent structure of the pilus-forming subunit of *S. pyogenes* revealed that intramolecular isopeptide bonds are used to stabilize the pilus building block, with one bond located in the vicinity of the sortase recognition site (Kang et al., 2007), although the precise role played by this specific arrangement in the pilus-forming process is still unclear.

In *S. pneumoniae* the entire macromolecular machinery responsible for pilus formation is encoded on a mobile pathogenicity islet (*rfrA*). Strains that lack the *rfrA* islet display less adherence to lung epithelial cells and are less virulent in murine models of

pneumonia and bacteremia, revealing that the pilus plays an important role in infection (Barocchi et al., 2006; Hava and Camilli, 2002). The *rlrA* islet harbors seven genes, three of which encode structural proteins (RrgA, RrgB, and RrgC) and three of which encode sortases (SrtC-1, SrtC-2, SrtC-3). Transmission electron microscopy images of infectious *S. pneumoniae* strains have revealed that the pilus shaft is composed of a RrgB polymer, with the minor pilins RrgA and RrgC associated to it in patches (Hilleringmann et al., 2008; LeMieux et al., 2006); in the case of RrgA, association has also been shown directly onto the cell wall (Barocchi et al., 2006). All six proteins are potentially linked to the cell membrane through a single, predicted transmembrane (TM) region. Notably, the three structural proteins carry an LPXTG-like motif immediately upstream of the TM region, which potentially enables them to be acted upon by the dedicated sortases. In addition, RrgB also displays a conserved Lys residue within a “pilin” motif that is present within most pilus-forming subunits from Gram-positive organisms. Thus, RrgB polymerization could proceed via a mechanism in which the LPXTG-like motif serves as a donor substrate, whereas a conserved Lys residue within a “pilin” motif potentially acts as the acceptor during a sortase-catalyzed transpeptidation reaction, in a process that would ensure stepwise, covalent incorporation of RrgB subunits into the growing fiber (Marraffini et al., 2006; Scott and Zahner, 2006; Telford et al., 2006). Participation of the ϵ -amino group of a conserved lysine residue in a nucleophilic pilus-forming reaction has now been shown for *Bacillus cereus* (Budzik et al., 2008) and *Corynebacterium diphtheriae* (Mandlik et al., 2008), suggesting that this might be a conserved mechanism in Gram-positive bacteria.

The equal number of structural and sortase-encoding genes in pneumococci has prompted the “one protein, one sortase” proposition (Hilleringmann et al., 2008; LeMieux et al., 2006), in which one sortase protein would be responsible for the polymerization of the RrgB fiber, whereas the two others could play roles in associating RrgA and RrgC (to RrgB and potentially to each other). In this work, we show that SrtC-1 is the main fiber-forming transpeptidase, and SrtC-3 has a modest capability of forming RrgB oligomers. Purified RrgB fibers produced in vitro by SrtC-1 activity are structural mimics of the pneumococcal pilus backbone. In addition, the high-resolution crystal structures of the two sortases that display in vitro reactivity toward RrgB, SrtC-1 and SrtC-3, reveal substrate-recognition regions that are covered by a lid that maintains the active sites in closed conformation, suggesting that pilus formation is regulated by a mechanism that involves activation of sortase-mediated catalysis. The identification of sequences corresponding to lid regions in sortases from a variety of pilated Gram-positive bacteria, as well as the important role played by pili in infection, suggest that pilus-forming sortases could represent tractable targets for antibiotic development for a wide spectrum of bacterial pathogens.

RESULTS

RrgB Fiber Can Be Assembled In Vitro by a Pilus-Polymerizing Transpeptidase

The pneumococcal *rlrA* pathogenicity islet encodes all of the genes required for formation of the pilus-forming machinery, namely backbone fiber (*rrgB*) and ancillary proteins (*rrgA*,

rrgC), as well as three sortases (SP0466, SP0467, SP0468, whose protein products are SrtC-1, SrtC-2, SrtC-3) (Barocchi et al., 2006; Hava and Camilli, 2002; LeMieux et al., 2008). To initiate our study of the pilus fiber formation process, we expressed the soluble forms of SrtC-1, SrtC-2, SrtC-3, and RrgB from the highly infectious *S. pneumoniae* serotype strain T4 (TIGR4) and tested the ability of each sortase to polymerize RrgB monomers. The presence of pili on the pneumococcal surface is characterized by a pattern of high-molecular-weight bands on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) upon analysis of surface preparations, which provide evidence of the covalent nature of the RrgB pilus backbone (Barocchi et al., 2006; Hilleringmann et al., 2008). Hence, the ability to form RrgB fibers in vitro was tested by searching for a pattern of high-molecular-weight bands on gradient SDS-PAGE upon incubation of RrgB monomers at 37°C with SrtC-1, SrtC-2, or SrtC-3. A ladder of bands with molecular weights above 100 kDa, which was recognized by antibodies raised against RrgB, could be generated when monomeric RrgB was incubated with SrtC-1 (Figure 1A, lane 2). In light of the fact that sortases catalyze transpeptidation reactions by employing a Cys nucleophile, we mutated the unique Cys residue within SrtC-1, Cys 193, into Ala, and tested this mutant protein for its ability to generate the RrgB high-molecular-weight ladder. Upon incubation with SrtC-1-Cys193Ala, RrgB remained monomeric, and no higher-molecular-weight bands could be detected (Figure 1A, lane 5), pointing to the fact that in vitro RrgB polymerization is a product of SrtC-1-related activity.

No evidence of RrgB polymerization could be identified upon incubation of RrgB with SrtC-2 (Figure 1A, lane 3); however, in the presence of SrtC-3, two faint bands could be reproducibly detected in the vicinity of the 117 kDa marker (Figure 1A, lane 4, indicated with stars), suggesting that SrtC-3 can catalyze the formation of higher-molecular-weight forms of RrgB, albeit much less efficiently than SrtC-1. These results suggest that although SrtC-1 is the most efficient transpeptidase in the RrgB polymerization process, SrtC-3 also displays a modest capability to form polymerized forms of RrgB.

To confirm these results, we deleted the gene encoding SrtC-1 (SP0466) from the *rlrA* islet in strain TIGR4. Total extracts of mutant TIGR4 as well as of the wild-type strain were treated with mutanolysin, loaded onto gradient polyacrylamide gels, and immunoblotted with antibodies specific for RrgB (Figure 1A, lanes 6–9). Wild-type TIGR4 presented high-molecular-weight bands characteristic of the presence of a polymerized pilus (Figure 1A, lanes 6 and 7), whereas the SP0466 deletion mutant presented a smaller quantity of fully polymerized pilus that was concomitant with an increase in monomeric RrgB. These results confirm that SrtC-1 plays a key role in the pilus polymerization process, and that its absence can be partly compensated by a “background” polymerization activity of other sortases.

In Vitro Reconstitution of the Pneumococcal Pilus Backbone

To obtain structural insight into the RrgB fibers prepared in vitro, we performed a large-scale incubation of SrtC-1 and RrgB and gel filtered the mix. The RrgB fiber eluted early in the gel-filtration experiment, whereas SrtC-1 and unpolymerized RrgB both eluted in volumes indicative of monomers (see Figure S1

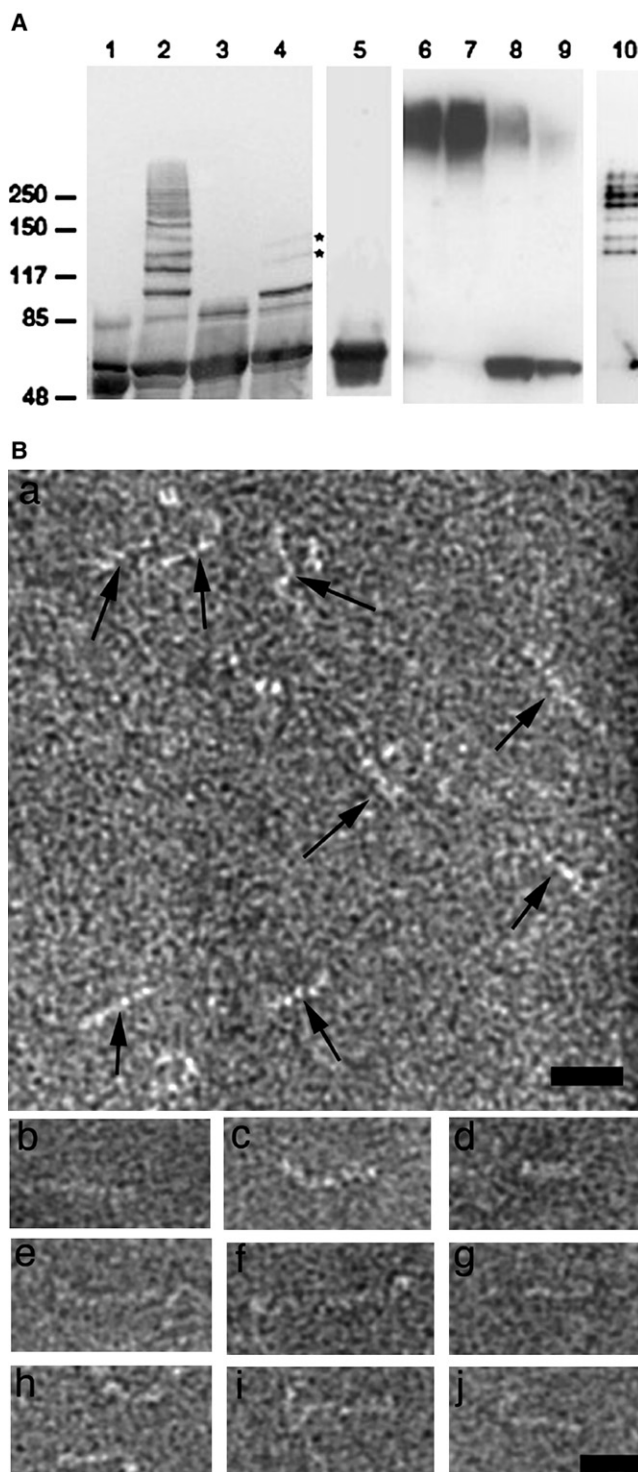


Figure 1. SrtC-1 Is a Pilus-Polymerizing Transpeptidase

(A) Western blots using polyacrylamide gradient gels developed with RrgB antiserum. Lanes: (1), monomeric RrgB; (2), RrgB + SrtC-1; (3), RrgB + SrtC-2; (4), RrgB + SrtC-3; (5) RrgB + SrtC-1Cys193Ala; (6) total extract from mutanolysed TIGR4; (7) supernatant from mutanolysed TIGR4; (8) total extract from mutanolysed TIGR4ΔSP0466; (9) supernatant from mutanolysed TIGR4ΔSP0466. Mutanolysed samples display pili that harbor not only polymerized RrgB, but also associated RrgA and RrgC, as well as peptidoglycan

available online). The high-molecular-weight sample (Figure 1, lane 10) was submitted to mass spectrometry (LCMS/MS) experimentation, which confirmed that the bands consisted of RrgB (Table S1); subsequently, it was analyzed by negative staining electron microscopy with a Philips CM12 microscope employing 2% sodium silicotungstate at pH 7.5. Electron micrographs show populations of fibers (Figure 1B, panel a) with roughly two distinct diameters: the most prevalent types resemble “beads on a string” and have diameters in the range of 3.5 nm, whereas the less common types of fibers have diameters of approximately 7.0 nm, thus being potentially representative of lateral associations of the thinner filaments (compare panels d and g in Figure 1B). Assuming that 1 image pixel = 1.17 Å, we estimate that the longest fibers have an approximate molecular weight of 480 kDa, whereas the shortest, of 140 kDa (see [Experimental Procedures](#) for precise calculations). In addition, several fibers display a “kink” (Figure 1B panels c, g, h, and i), giving evidence of flexibility. Notably, electron microscopy analyses of pili on the surface of *S. pneumoniae* have shown them to be flexible structures formed by 3.5-nm-thick protofilaments that intersect laterally to form pili with diameters of 6.8 nm or, in areas of more relaxed intersection, 9.5 nm (Hilleringmann et al., 2008). Our observations suggest that the fibers formed in vitro through the covalent association of RrgB monomers by SrtC-1 are structural mimics of the 3.5 nm and 6.8 nm filaments formed on the pneumococcal surface.

A Closed Substrate-Recognition Channel within Pilus-Forming Sortases

To gain insight into the catalytic aspects of the pilus-forming machinery at an atomic level, we crystallized and solved the high-resolution structures of SrtC-1 and SrtC-3. Both enzymes elute in volumes indicative of monomers in gel filtration, but crystallize as dimers in the asymmetric unit of a P2₁2₁2₁ (SrtC-1) or P2₃ (SrtC-3) cell. The structure of SrtC-1 was solved by performing a 4 wavelength MAD experiment on a selenomethionylated SrtC-1 crystal at the ESRF synchrotron in Grenoble; this model was subsequently used to phase native data to 1.24 Å. SrtC-3 data were collected to 2.2 Å on an in-house X-ray generator and the structure was solved by employing the refined structure of SrtC-1 (with modifications; see [Experimental Procedure](#)) as a search model in a molecular replacement experiment. Data collection and refinement statistics can be found in [Tables 1 and 2](#).

SrtC-1 (residues 21–214; [Figure 2A](#)) is composed of eight central, mostly anti-parallel β-strands that associate to form a β barrel that is surrounded by three major α helices and three 3₁₀ helices. Helices α1, α2, and α3 form the roof of the SrtC-1 β-barrel, with the region that includes α3 forming a “lid” that covers the active site (pink in [Figures 2A and 2C](#)). SrtC-3 (residues 39–241; [Figures 2B and 2C](#)), despite sharing 26.4% sequence identity with SrtC-1, displays a very similar fold to the latter enzyme (rms deviation of 1.3 Å over 138 Cα atoms), with

fragments. (10) Gel-filtered sample analyzed by electron microscopy and LCMS/MS.

(B) Top, negative staining electron microscopy images of the RrgB fibers generated by SrtC-1 activity. Black arrows point to fibers whose diameters are in the range of 3.5 nm or 7.0 nm. Bottom, a gallery of typical fibers seen on grids. Scale bar, 20 × 5 nm.

Table 1. Data Collection and Phasing Statistics for SrtC-1 and SrtC-3

Data Collection						
Data set	Peak	Inflection Point	High Remote	Low Remote	SrtC-1 Native	SrtC-3 Native
Wavelength (Å)	0.98055	0.980750	0.982450	0.978550	0.93100	1.5418
Space Group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P23
a (Å)	68.286	68.619	68.757	68.662	68.091	122.638
b (Å)	70.070	70.366	70.469	70.337	70.189	122.638
c (Å)	86.187	86.647	86.807	86.673	86.542	122.638
Resolution (Å)	1.73 (1.83–1.73)	1.86 (1.98–1.86)	1.86 (1.98–1.86)	1.86 (1.98–1.86)	1.24 (1.32–1.24)	2.14 (2.27–2.14)
No. obs/unique ref.	297168/80448	240375/65317	240388/65393	236792/64343	1493780/114617	205227/30999
Completeness (%)	96.3 (88.0)	96.7 (89.3)	95.6 (84.8)	95.1 (83.5)	98.3 (93.9)	90.7 (75.8)
R _{sym} (last shell)	5.9 (37.4)	5.7 (44.6)	6.1 (63.6)	7.5 (83.4)	4.7 (50.7)	10.0 (51.0)
I/σ(I) (last shell)	13.4 (3.6)	13.7 (3.2)	12.1 (2.3)	10.1 (1.8)	33.5 (5.1)	15.3 (4.4)
SrtC-1 Phasing						
SHARP FOM (acentric/centric)	0.43849/0.14854	0.43849/0.14854	0.43849/0.14854	0.43849/0.14854		
Phasing power ISO (acentric/centric)		0.253/0.208	0.844/0.879	0.170/0.160		
Phasing power ANO	1.807	1.110	0.102	0.643		

the major difference being the presence of an additional α -helix at the C terminus of SrtC-3. SrtC-3 also displays a “lid” over the active site region; however, in both SrtC-1 and SrtC-3, the lid regions display the highest temperature factors in the structures, suggesting that they could be flexible or mobile elements. Analysis of the SrtC-3 structure reveals that residues 68–71 and 78–82, located on the SrtC-3 lid, display B-factors of 47.4 Å² and 44.8 Å², respectively; these values are higher than for those of the intervening region (residues 73–76, B-factor = 20.2 Å²) as well as for the overall structure B-factor (23.8 Å²; Figure S2). These observations suggest these residues could play the roles of hinges in a potential lid opening mechanism (see below). The central section of SrtC-1 and SrtC-3, which encompasses the β -barrel, is reminiscent of that seen in the structure of sortases involved in association of proteins to the peptidoglycan (SrtA) or iron-heme transport (Ilangovan et al., 2001; Zhang et al., 2004). However, the arrangement of α helices in the surrounding regions is distinct, and the lid seen in the structures of the pilus-

forming sortases related here is absent from all other sortases whose structures have been solved to date (Figure S3). In addition, biochemical studies performed on SrtA from *S. aureus* have revealed that its β 6/ β 7 loop is a potential substrate specificity determinant (Bentley et al., 2007), and it is affected by the presence of Ca²⁺ (Naik et al., 2006). Although the specificity of the β 6/ β 7 region of SrtC-1/SrtC-3 was not studied here, we did not observe any effect of Ca²⁺ on RrgB fiber-forming activity (not shown).

All sortases recognize the LPXTG motif of a target protein, employ their catalytic cysteine in the nucleophilic attack of the Thr-Gly peptidic bond, and subsequently catalyze a transpeptidation reaction by employing the free amino group from an acceptor substrate (Ton-That et al., 2002; Zong et al., 2004a). The SrtC-1 active site includes a Cys-His-Arg triad that has been implicated in catalysis in other surface-attachment sortases: Cys 193, at the C terminus of β 7; Arg 202, at the N-terminus of β 8; and His 131, within the loop that follows β 4 (Figure 3A). The side chain of Cys 193 is present in two conformations (in both monomers of the asymmetric unit) and was modeled with 0.5 occupancy in each case. In one of the conformations, it points away from the center of the active site and makes a unique, 3.1 Å hydrogen bond with the O δ 1 atom of Asn 199, itself located on the loop between β 7 and β 8. In the second conformation, the S γ moiety of Cys 193 points directly into the active site, and makes a weak 3.4 Å hydrogen bond with the NH1 atom of Arg 202. Notably, as is the case for other sortases whose structures have been solved to date (Ilangovan et al., 2001; Zhang et al., 2004; Zong et al., 2004b), the nucleophilic cysteine does not interact directly with His 131 (the closest distance is 5.4 Å). However, it is of note that the lid contributes two residues that are anchored within the active site: Asp 58 and Trp 60. Asp 58 guarantees that the Arg 202 faces the Cys nucleophile by interacting with both NH2 and N ϵ atoms, whereas Trp 60, in addition to considerably increasing the hydrophobic nature of the cleft, interacts with the N δ 1 moiety of His 131 through its N ϵ 1 atom.

Table 2. Refinement Statistics for SrtC-1 and SrtC-3

Refinement	SrtC-1	SrtC-3
Resolution (Å)	1.24	2.14
R _{work} (%)	17.31	18.05
R _{free} (%)	19.31	24.14
No. of protein atoms	3076	3340
No. of solvent atoms	471	505
No. of glycerol molecule	1	0
RMS deviation, bond lengths (Å)	0.007	0.028
RMS deviation, bond angles (°)	1.250	2.296
Mean B-factor (Å ²)	16.66	37.70
Res. most favored/allowed regions of Ramachandran (%)	100	100

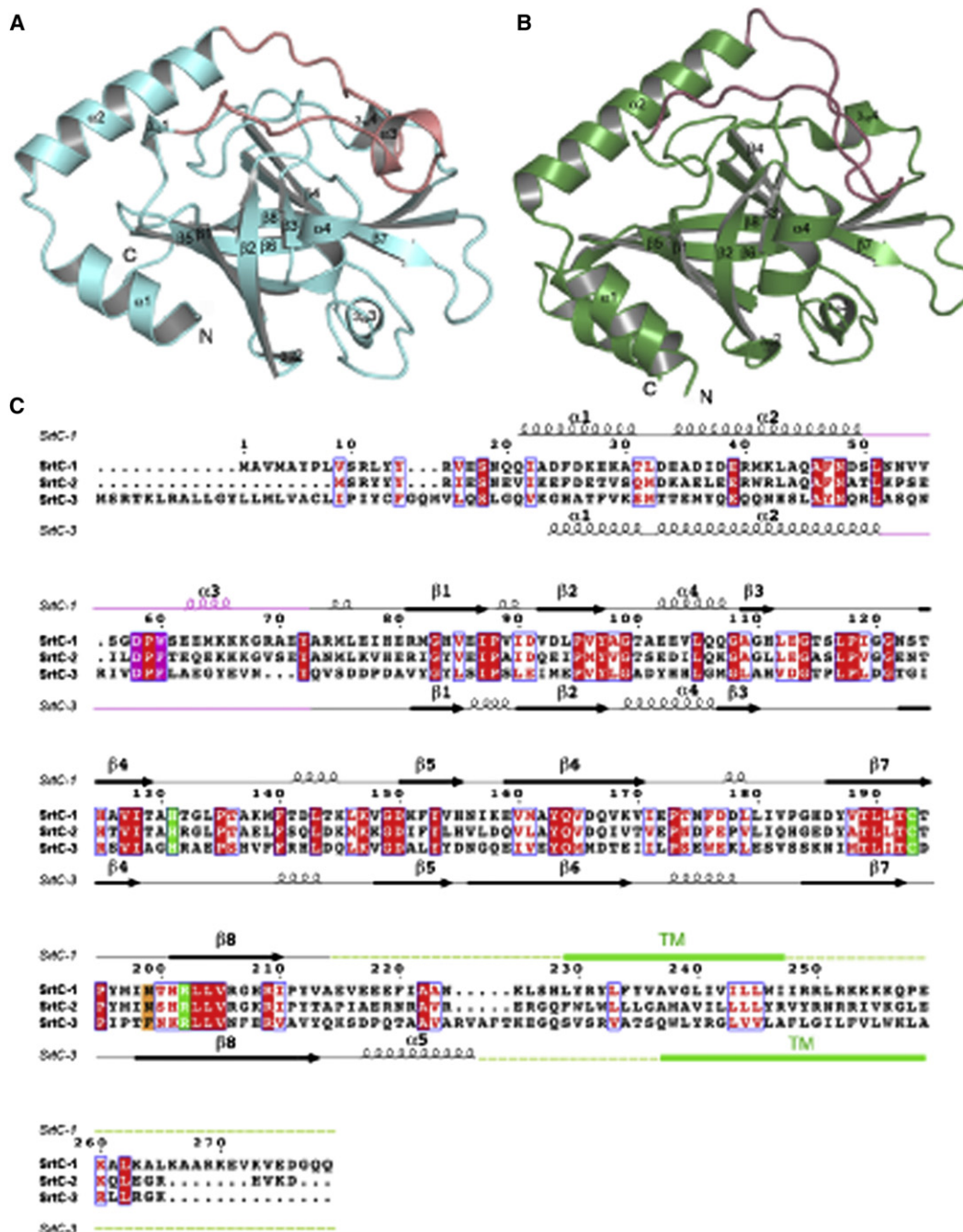


Figure 2. Sortases of the Pilus-Formation Machinery Display Similar Folds

(A and B) SrtC-1 (A) and SrtC-3 (B) fold into β -barrels surrounded by helices; their active sites, centered on the region surrounding the C terminus of $\beta 7$ and the N terminus of $\beta 8$, are covered by a lid (shown in pink).

(C) Structure-based sequence alignment of the three pilus-forming sortases of *S. pneumoniae*. Identical residues are shown with a red background, whereas similar residues are shown in red and highlighted with blue boxes. Residues located within the active site cleft of SrtC-1/SrtC-3 are highlighted in green, whereas the Asn residue on $\beta 7/\beta 8$ that contacts the nucleophilic Cys in SrtC-1 is shown with an orange background. The lid region is shown in magenta. The TM (transmembrane) region was predicted with the SMART server (www.smart.embl-heidelberg.de).

The active site of SrtC-3 shows clear similarities with that of SrtC-1; Cys 206, Arg 215, and His 144 occupy analogous positions to their counterparts in SrtC-1. In this case, however, a single conformation of the Cys 206 side chain was detected, and it

points within the active site, toward Arg 215 (Figure 3B). This difference could be caused by the fact that Asn 199 of SrtC-1 is has been substituted by a Phe side chain in SrtC-3, and thus the Cys-Asn polar interaction observed in SrtC-1 is not possible

here. In addition, the SrtC-3 lid gives evidence of being more flexible than that of SrtC-1, with only the region encompassing the anchor residues being clearly traceable in the electron density map. These residues include Asp 73, which, as is the case in SrtC-1, interacts directly with Arg 215, and Phe 75, which points snugly into the pocket and guarantees its hydrophobic nature (much like Trp 60 in SrtC-1). It is of note that in both structures, the Asp residue located at the edge of the lid locks the Arg side chain in place, positioning it in close proximity to the nucleophilic Cys. Interestingly, the structures of surface-attachment sortases have revealed active sites that are not covered by lids and are much more solvent-exposed than those observed here (Figure 3C). In addition, these exposed active clefts display catalytic residues whose side-chain positions are not constrained; in SrtA from *S. aureus*, even in the presence of substrate, the side chain of the catalytic Arg residue points away from the nucleophilic Cys (Zong et al., 2004a). This observation underlines the possibility that the role of the interaction between the catalytic Arg and the Asp side chain located on the SrtC-1/SrtC-3 lid is not only to “lock” the lid in place in the absence of substrate, but also to guarantee that the Arg side chain will be appropriately positioned, potentially for oxyanion transition state stabilization (Frankel et al., 2007), once the lid is removed and substrate is bound. Thus, pilus-forming sortases present tightly knit active sites whose access is restricted by the presence of a potentially flexible lid, and whose residues are pre-positioned for catalysis.

DISCUSSION

Although pili have been identified on the surface of approximately 30% of the strains of the major human pathogen *S. pneumoniae* and shown to be involved in pathogenicity, little is known about the complex macromolecular machinery that regulates their biogenesis. In this work, we have identified that formation of the pilus backbone in the pneumococcus, known to be composed of the covalently linked subunits of RrgB, is catalyzed by a pilus-polymerizing enzyme, SrtC-1, the first sortase within the *rlrA* pathogenicity islet. We show that SrtC-1 is able to polymerize monomeric RrgB in vitro, whereas the third sortase encoded by the islet, SrtC-3, displays a modest capacity to associate RrgB monomers. A pneumococcal strain in which the SP0466 gene was knocked out displays a clearly diminished capacity of producing polymerized pili on its surface as well as an increased amount of RrgB monomers associated to the bacterial cell wall. Interestingly, a small amount of pili can still be detected on the surface of the TIGR4ΔSP0466 strain, suggesting that the absence of this pilus-polymerizing transpeptidase can be partly compensated by another sortase. These observations are corroborated by those of LeMieux and coworkers (LeMieux et al., 2008), who show that pneumococcal strains lacking individual pilus-forming sortases can still express pili, although the amount of surface-associated pili formed was not explored in their study.

In addition, we show that the RrgB fibers can be reconstituted in vitro from purified RrgB monomers and SrtC-1. In this sense, SrtC-1 acts as a catalyst to promote the covalent association of RrgB monomers independent of cellular energy. Pilus fibers produced in vitro and visualized in this work by electron microscopy

have approximately the same diameters as the single-coiled (3.5 nm) or double-coiled coil (6.8 nm) protofilaments identified and measured on the surface of pneumococci (Hilleringmann et al., 2008). Nevertheless, the assemblies produced in vitro (average length 35 nm) are shorter than the longest pili yet visualized in vivo, which have been reported as having lengths of up to 1 μ m (Hilleringmann et al., 2008). It is of interest, however, that surface-associated pili are often heterogeneous in length, with short forms also being visible in several instances (Barocchi et al., 2006; Hilleringmann et al., 2008; Nallapareddy et al., 2006; Rosini et al., 2006; Weerkamp et al., 1986). It is thus possible that the structures isolated and visualized in this work could represent the minimal building blocks of the pilus fiber. However, it is also possible that the RrgB fibers formed in vitro are shorter than their wild-type counterparts because SrtC-1 polymerizes RrgB inefficiently (as shown by the finite number of high-molecular-weight bands on Figure 1A, lane 2). This could indicate that efficient fiber formation requires the concomitant participation of other members of the pilus-forming machinery, such as the other sortases. However, when this possibility was tested in vitro by incubation of RrgB with SrtC-1, SrtC-2, and SrtC-3 in different combinations, the pattern of formed fibers was identical to that observed with SrtC-1 only (not shown). It is thus conceivable that, in vivo, fiber formation is a well-orchestrated event that requires that the entire pilus-forming macromolecular machinery be associated to the bacterial outer membrane in order to optimize partner recognition events and catalysis, an organization that cannot easily be emulated in vitro.

Our results also reflect the potential requirement that SrtC-1 be maintained in an “activated” state through the entire fiber-formation process. Evidence for this possibility comes from the mapping of conserved/identical residues between sortases on the surface of SrtC-1, and comparison with the surface diagram of a well-studied sortase, SrtA from *S. aureus* (Figures 2C and 4). All sortase isoforms studied to date possess a highly conserved Thr-Leu-X-Thr motif that precedes the Cys nucleophile, and is located along β 7 (shown in orange and red in Figure 4). Modifications within this sequence in SrtA suggest that it plays a role in substrate recognition (Frankel et al., 2007), a conclusion that is supported by the crystal structure of SrtA in complex with the LPETG peptide, which reveals that the substrate is located in proximity to this sequence (Zong et al., 2004a) (Figure 4B). Both SrtC-1 and SrtC-3 also possess this highly conserved region within an elongated cleft; however, in the pilus-forming sortases, this zone is covered by the lid, whose function also involves providing an anchor for the active site (shown in dark blue). These observations suggest that, in SrtC-1/SrtC-3, a closed lid (as seen in the structure) potentially prevents recognition of LPXTG-like sequences, necessitating a mechanism of lid opening in order for catalysis to occur. This process could be initiated by an initial recognition between the Rrg-substrate and the closed sortase; only binding of the appropriate substrate would lead to lid opening and enzyme activation. This model would be in agreement with the one recently put forward by Mandlik and coworkers (Mandlik et al., 2008), who suggest that, in order for proper pilus assembly to occur, two distinct sortases must act sequentially: the pilus-forming and the housekeeping enzymes, the latter of which will recognize lipid II as a substrate and anchor the formed polymers to the cell wall. Because, as

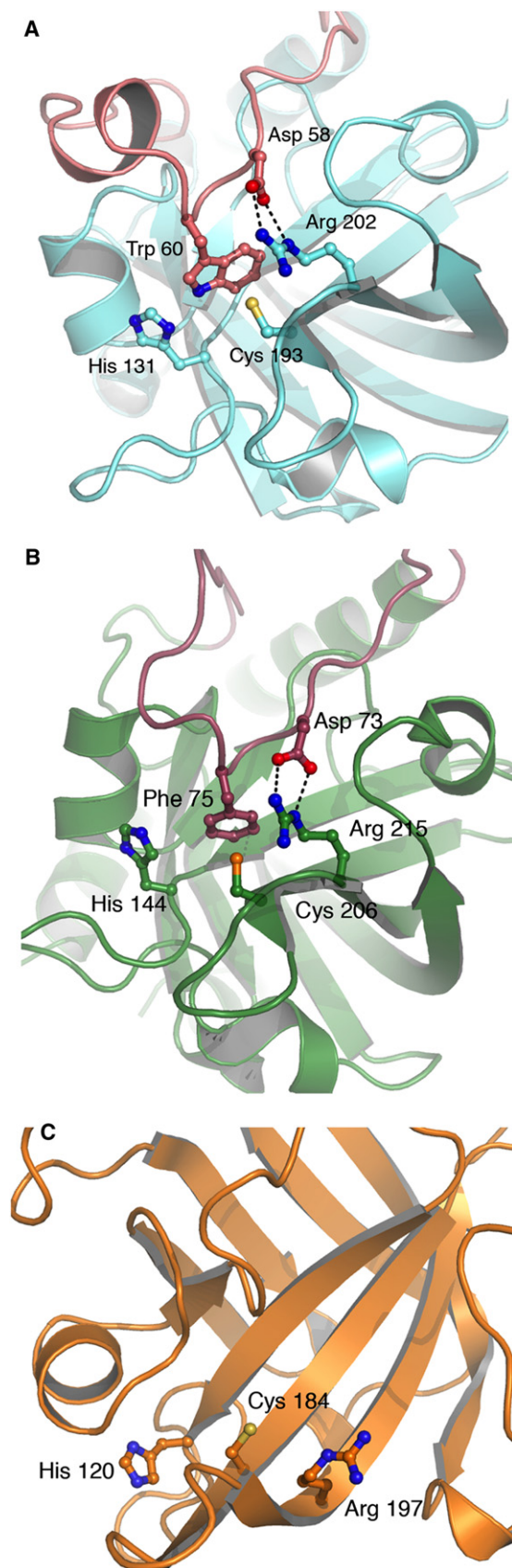


Figure 3. The Active Site of Pilus-Forming Sortases Is Anchored by a Lid

(A) The nucleophilic Cys 193 in SrtC-1 displays two conformations, but only one is shown for clarity.

(B) The active site of SrtC-3, like SrtC-1, displays a catalytic Arg residue whose side chain is stably anchored by an Asp residue on the lid “anchor” region. In both active sites, the conserved His residue is approximately 5.4 Å away from the Cys nucleophile.

(C) The active site of SrtA from *S. aureus* (Protein Data Bank code 1T2P) is mostly solvent-exposed and does not harbor a lid.

shown in this work, the active sites of both sortase types are very similar, encapsulation of the catalytic cleft of the pilus-forming enzyme by a lid that can only be opened by the appropriate pilin subunit would ensure that lipid II is not recognized, thus guaranteeing proper orchestration of polymerization and cell wall attachment steps.

All three structural proteins that compose the pneumococcal pilus carry LPXTG-like motifs on their C termini, albeit with slightly distinct sequences (RrgA: YPRTG; RrgB: IPQTG; RrgC: VPDTG), which in principle are the substrates of a comparable transpeptidation reaction. Their exquisite specificity was illustrated by the elegant work performed by LeMieux and coworkers (LeMieux et al., 2006), who showed that exchange of the YPRTG motif of RrgA into the analogous motifs of RrgB or RrgC abrogated association of RrgA onto the RrgB fiber. Hence, slight modifications within the LPXTG-like sequences might have profound consequences for catalysis. To gain insight into the question of specificity of pilus-forming sortases, we compared surface potentials of all three enzymes after generating a model of SrtC-2 based on the structure of SrtC-1 (with which it shares 55.7% sequence identity; Figure 2C). This analysis (Figure 5) reveals clear differences within the active sites of SrtC-1, SrtC-2, and SrtC-3, as well as the regions that surround them. Most striking is the vast acidic region on the surface of SrtC-3, which contrasts sharply with the much smaller acidic region on the surfaces of SrtC-1 and SrtC-2, centered on Glu 103 and Glu 94, respectively; these differences are located on the closed lid regions. In addition, the catalytic clefts themselves also show differences, with that of SrtC-1 being more basic than that of SrtC-2, whereas that of SrtC-3 displays a clear negative charge. Thus, sortase specificity and differential recognition of RrgA, RrgB, and/or RrgC potentially involve interactions not only within the active site, but also with the surrounding region, which includes the flexible lid, as suggested above. It is of interest, thus, that our experiments show that both SrtC-1 and SrtC-3 can associate RrgB monomers, confirming that a certain level of “cross-specificity” exists among pilus-forming sortases. In light of our results, we suggest that this might arise as a consequence of the considerable flexibility observed for the lid region, coupled to the conservation of active site residues and the nature of the transpeptidation reaction catalyzed. These observations give the sortase lid region a key role in the processes of pilin substrate recognition and initiation of catalysis.

To verify if these observations could be extended to the pilus biogenesis process in other Gram-positive organisms, we performed a BLAST analysis using the pneumococcal SrtC-1 sequence and searched for the telltale lid anchor residues

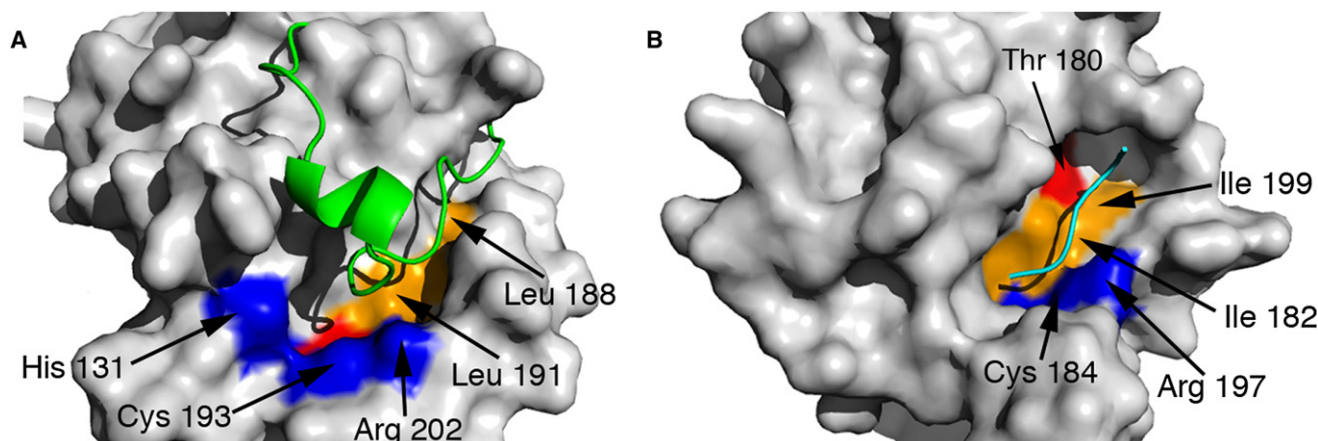


Figure 4. The Substrate-Binding Cleft of Sortases Is Conserved

(A) Surface of SrtC-1, onto which residues that are conserved (orange) and identical (red) among the three pilus-forming sortases and have been mapped. The analogous residues were mapped onto the surface of *S. aureus* SrtA (B). This analysis identifies that the lid of pilus-forming sortases (shown in green) is positioned analogously to the LPETG substrate peptide (shown in cyan) in the structure of SrtA, suggesting that the substrate-recognition cleft in pilus-forming sortases is sterically blocked by the lid. Active site residues are shown in dark blue.

(Asp-Pro-Hyd) located approximately 125 amino acids N terminal to the nucleophilic Cys (magenta in Figure 2C). Interestingly, we were able to identify this sequence in pilus-forming sortases from a variety of other pilated species, including *Streptococcus agalactiae*, *Streptococcus pyogenes*, *Bacillus cereus*, *Enterococcus faecium*, *Actinomyces naeslundii*, and the well-studied *Corynebacterium diphtheriae* (Figure S4). These observations suggest that Gram-positive organisms could employ a universal mechanism for pilus biogenesis, where sortases encoded by the pilus-related pathogenicity islets harbor encapsulated active sites, thus necessitating substrate-induced activation for pilus formation.

The targeting of pathways essential for bacterial survival by antibiotics has a strong link to the development of drug resistance. Sortases are not essential for bacterial survival, but they allow bacteria to express surface-associated proteins that play important roles in virulence. Developing molecules that challenge a bacterium's capability to successfully employ its pathogenicity

arsenal during infection could be an alternative approach toward antibiotic development (Walsh, 2003). The absolute requirement for sortases for pilus formation in Gram-positive pathogens, coupled to active site features that clearly distinguish them from classic sortases and to the fact that they are surface-exposed and thus potentially ligand-accessible, suggests that these enzymes are tractable, potential targets for novel antibiotic development for a variety of Gram-positive species.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification of RrgB, SrtC-1, SrtC-2, and SrtC-3

The regions of the *S. pneumoniae* TIGR4 coding for amino acids 30–633 (RrgB), 17–228 (SrtC-1), 9–220 (SrtC-2), and 32–254 (SrtC-3) were amplified using conventional PCR methodologies and cloned into vector pLIM (Protein Expert, Grenoble). Resulting vectors were transformed into *E. coli* RIL cells (Invitrogen). In all four cases, protein expression was induced in Terrific Broth with 1 mM IPTG at 37°C during 3 hr. Cells were harvested by centrifugation

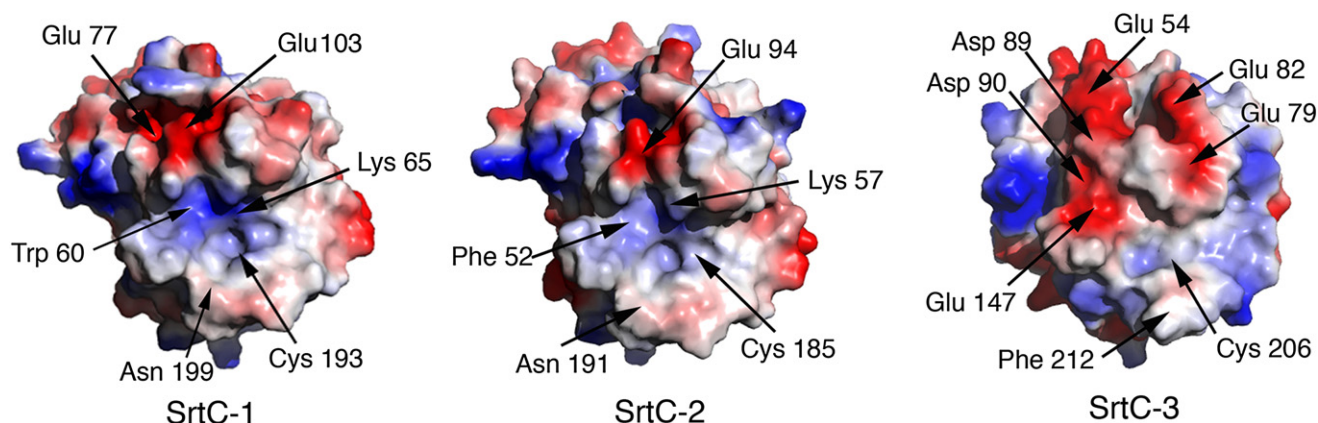


Figure 5. Surface Potential Representations of All Three Pilus-Forming Sortases, with a Direct View onto the Active Site

The structure of SrtC-2 was generated by building a model based on the structure of SrtC-1, with which it shares 55.7% sequence identity. Striking features include the highly basic character of the SrtC-1 cleft and the vast acidic region in the vicinity of the SrtC-3 active site (same orientation as in Figure 3).

and lysed by sonication in 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 20 mM imidazole, 1 mM PMSF, 0.1 μ M aprotinin, and 1 μ M pepstatin. The lysates were clarified by centrifugation and applied onto HisTrapTMHP columns (Amersham Biosciences) pre-equilibrated in lysis buffer. Protein elution was performed using linear gradients of 20–500 mM imidazole over 20 ml. Protein-containing fractions were dialyzed 2 hr into 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, and 20 mM imidazole at 4°C and subsequently incubated with 1:10 (w/w) Tev protease overnight at 15°C. The proteins were reloaded onto HisTrapTMHP columns and the eluted, cleaved products were loaded onto HiLoad 16/60 Superdex 200 (Amersham Biosciences) in 50 mM HEPES (pH 7.5) and 150 mM NaCl. Pooled, concentrated fractions were employed in crystallization trials and RrgB fiber formation experiments as described below.

Selenomethionylated SrtC-1 was overexpressed in *E. coli* RIL cells in minimal medium supplemented with thiamine (0.2 mg/ml), leucine (50 mg/l), valine (50 mg/l), isoleucine (50 mg/l), lysine (100 mg/l), phenylalanine (100 mg/l), threonine (100 mg/l), and selenomethionine (60 mg/l). Expression and purification was performed as described above for the native SrtC-1.

In Vitro RrgB Fiber Formation and Electron Microscopy Analyses

Recombinant, purified RrgB and SrtC-1 were mixed at a ratio of 1:2 (w/w) and incubated overnight at 37°C. RrgB was also incubated with SrtC-2 and SrtC-3 at the same ratio and under the same conditions, in a total volume of 40 μ l. Reactions were stopped by boiling for 10 min in the presence of Coomassie blue dye and 1% SDS. Aliquots of the mixtures were subsequently analyzed on 4%–12% SDS-PAGE gradients and then subjected to immunoblotting using a 1:5000 dilution of mouse antibodies raised against RrgB. After washing, membranes were incubated with an anti-mouse antibody conjugated to horseradish peroxidase at a 1:10,000 dilution. An immunoreactive signal was detected by chemiluminescence (ECL, Biorad).

Samples destined to electron microscopy analyses were prepared as described above and subsequently loaded onto a Superdex 200 gel filtration column (GE Healthcare). All eluted fractions were tested by immunoblotting with anti-RrgB antibodies, and those that corresponded to the high-molecular-weight forms (Figure 1A) were submitted to negative staining electron microscopy analyses.

Four microliters of the protein sample at approximately 0.05 mg/ml were adsorbed onto the clean face of a carbon film on a mica sheet (carbon/mica interface) and negatively stained with 2% (w/v) neutral sodium silicotungstate. Micrographs were taken under low-dose conditions with a CM12 LaB6 electron microscope working at 120 kV and with a nominal magnification of 60,000X. Electron micrographs were digitized using a Zeiss scanner with a pixel size of 7 nm (1.17 \AA at the sample scale).

Volumes of individual protein fibers were calculated by assuming that 1 pixel = 1.17 \AA (and thus $50 \text{ \AA} = 43$ pixels), and that the average density of a protein is 0.86 Da/\AA^3 . The calculated volume will thus correspond to the length of each individual fiber (L) multiplied by the surface of the section ($\pi \times r \times r$, where r represents the fiber at half-height). Thus:

for panel (a), $290 \text{ \AA} \times 50 \text{ \AA}$; $L = 290 \text{ \AA}$; $r = 25 \text{ \AA} \rightarrow 480 \text{ kDa}$

for panel (d), $140 \text{ \AA} \times 54 \text{ \AA}$; $L = 140 \text{ \AA}$; $r = 27 \text{ \AA} \rightarrow 275 \text{ kDa}$

for panel (h, bottom), $120 \text{ \AA} \times 42 \text{ \AA}$; $L = 120 \text{ \AA}$; $r = 21 \text{ \AA} \rightarrow 140 \text{ kDa}$

The fibers shown in the EM images thus have masses that range from 140 kDa to approximately 480 kDa.

Construction and Analysis of the SP0466 Deletion Mutant

The SP0466-inactivated TIGR4 strain was constructed by PCR insertion of the *cat* cassette directly into the genome. The PCR product was transformed into the wild-type TIGR4 strain, as previously described (Bricker and Camilli, 1999) and then plated onto Columbia blood agar plates containing chloramphenicol (4 μ g/ml). Transcription of the downstream genes (SP0467, SP0468) was verified by analysis of the total RNA of the mutant, demonstrating no polar effect as a consequence of the SP0466 deletion (data not shown). Wild-type TIGR4 and deletion strains were grown in Todd-Hewitt broth (BD) supplemented by 0.5% yeast extract (Sigma) at 37°C, 5% CO_2 , until an $\text{OD}_{600\text{nm}}$ of 0.45 A.U. Cultures were centrifuged and the bacterial pellet, after being washed with PBS,

was treated with 200 U/ml mutanolysin (Sigma), 1 mg/ml lysozyme (Sigma), and a protease inhibitor cocktail (Complete, Roche) for 3 hr at 37°C. Cellular debris were removed by centrifugation at 11 000 rpm for 15 min.

Thirty microliters of each sample was mixed with XT Sample Buffer and XT Reducing Agent (Bio-Rad), boiled at 100°C for 10 min and loaded onto 4%–12% Criterion XT Precast gels (Bio-Rad). Gels were run for approximately 3 hr and subsequently electrotransferred in Trans-Blot Transfer Medium (Bio-Rad). Incubation times of 1 hr were successively performed using anti-RrgB polyclonal mouse antibodies (diluted 1:5000) and anti-mouse HRP conjugate antibodies (Sigma, diluted 1:120,000) before detection with a chemiluminescent substrate (Pierce).

Crystallization, Data Collection, Structure Solution, and SrtC-2 Modeling

Crystals of native SrtC-1 were obtained by the hanging-drop vapor diffusion method in 100 mM Tris-HCl (pH 8.5), 200 mM MgCl_2 hexahydrate, and 30% w/v polyethylene glycol 4000 at 20°C. Crystals were cryoprotected by brief incubation in mother liquor containing increasing concentrations of glycerol (up to 15%), and subsequently were flash-cooled in the nitrogen stream. A first native data set (1.24 \AA) was collected at the European Synchrotron Radiation Facility (ESRF) beamline ID14-EH3 (Grenoble, France). A four-wavelength MAD experiment was performed on the Se edge on the ESRF ID23-EH1 beamline.

Diffraction images were indexed and scaled with XDS (Kabsch, 1993) and merged with the CCP4 6.0.2 program suite (CCP4, 1994). Identification of selenium atom positions, refinement, and phasing were performed with AutoSHARP (Bricogne et al., 2003; de la Fortelle and Bricogne, 1997) and automatic model building was performed with ARP/wARP 7.0.1 (Perrakis et al., 1999). PHASER (Storoni et al., 2004) was subsequently used to perform molecular replacement using the model generated by the MAD experiment in order to phase data collected to 1.24 \AA , and the model was automatically completed and refined using LaFite 2.6 (Yao et al., 2006). Cycles of manual model building were performed with COOT 0.4.1 (Emsley and Cowtan, 2004), whereas cycles of restrained refinement employing TLS motion determination (Painter and Merritt, 2006) were performed with REFMAC 5.4 (Murshudov et al., 1997) as implemented in the CCP4 program suite. Data collection and refinement statistics are shown in Tables 1 and 2. Residues 20–213 of SrtC-1 are included in the final model. There is an rmsd value of 0.211 \AA between the two monomers that form the asymmetric unit dimer.

SrtC-3 crystals were obtained in 100 mM MES (pH 5.0), 0.5 M $(\text{NH}_4)_2\text{SO}_4$, and 1.0 M LiSO_4 , and cryocooled directly on the nitrogen stream. A native data set was collected to 2.14 \AA using an in-house Rigaku rotating anode generator equipped with osmic multilayer mirrors and a R-Axis IV²⁺ detector. The structure was solved by molecular replacement using the program PHASER (Storoni et al., 2004) and employing a search model consisting of the core of SrtC-1 (residues 82–209) in which all amino acids that differed between SrtC-1/SrtC-3 were replaced by alanines. Manual model building and refinement were performed as for SrtC-1, as described above. For both structures, stereochemical verification was performed by PROCHECK (Laskowski et al., 1993) and secondary structure assignment was performed by DSSP (Kabsch and Sander, 1983) and STRIDE (Heinig and Frishman, 2004). The final SrtC-3 model includes residues 35–242, and there is an rmsd of 0.045 \AA between the two monomers that generate the dimer seen in the asymmetric unit. Data collection and structure refinement statistics are shown in Tables 1 and 2. The SrtC-2 model was generated through homology modeling based on the structure of SrtC-1 followed by refinement, using the Modeler 9.3 program (Sali and Blundell, 1993) as implemented in CCP4. Figures were generated with PyMol (<http://www.pymol.org>).

ACCESSION NUMBERS

The coordinates of SrtC-1 and SrtC-3 have been deposited in the Protein Data Bank under the ID codes 2w1j and 2w1k, respectively.

SUPPLEMENTAL DATA

Supplemental Data include four figures, one table, Supplemental Methods, and Supplemental References and can be found with this article online at [http://www.cell.com/structure/supplemental/S0969-2126\(08\)00385-7](http://www.cell.com/structure/supplemental/S0969-2126(08)00385-7).

ACKNOWLEDGMENTS

The authors thank Xavier Vernede (LCCP, IBS) and the ESRF ID14-EH3 beam-line staff for help with data collection, the IBS mass spectroscopy facility (LSMP, IBS) for analyses, Alexandra Kraut (Lab. d'Etude de la Dynamique des Protéomes, iRTSV Grenoble) for LCMS/MS experiments, Jean-Pierre Andrieu (LBM, IBS) for N-terminal sequencing, Claire Durmort (LIM, IBS) for help with RNA analysis, Véronique Frachet and the RioBioMol facility (LIM, IBS) for gene cloning, and J. Marquez and the HTX Lab team (Partnership for Structural Biology) for access to and help with high-throughput crystallization. This work was partly supported by Young Investigator (ANR Jeunes Chercheurs) grants to A.M.D.G. (n° 05-JCJC-0049) and G.S. (n° 06-JCJC-0126), as well as European Commission grant LSHM-CT-2004-512138 (to A.D. and T.V.).

Received: July 11, 2008

Revised: October 3, 2008

Accepted: October 9, 2008

Published: December 9, 2008

REFERENCES

- Barocchi, M.A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernebro, J., Moschioni, M., Masignani, V., et al. (2006). A pneumococcal pilus influences virulence and host inflammatory responses. *Proc. Natl. Acad. Sci. USA* 103, 2857–2862.
- Bentley, M.L., Gaweska, H., Klelec, J.M., and McCafferty, D.G. (2007). Engineering the substrate specificity of *Staphylococcus aureus* sortase A. *J. Biol. Chem.* 282, 6571–6581.
- Bogaert, D., Hermans, P.W., Adrian, P.V., Rumke, H.C., and De Groot, R. (2004). Pneumococcal vaccines: an update on current strategies. *Vaccine* 22, 2209–2220.
- Bricker, A.L., and Camilli, A. (1999). Transformation of a type 4 encapsulated strain of *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.* 172, 131–135.
- Bricogne, G., Vonnrhein, C., Flensburg, C., Schiltz, M., and Paciorek, W. (2003). Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D Biol. Crystallogr.* 59, 2023–2030.
- Bronzwaer, S., Cars, O., Buchholz, U., Molstad, S., Goettsch, W., Veldhuijzen, I.K., Kool, J.L., Sprenger, M.J.W., and Degener, J.E. (2002). A European study on the relationship between antimicrobial use and antimicrobial resistance. *Emerg. Infect. Dis.* 8, 278–282.
- Budzik, J.M., Marraffini, L.A., Souda, P., Whitelegge, J.P., Faull, K.F., and Schneewind, O. (2008). Amide bonds assemble pili on the surface of bacilli. *Proc. Natl. Acad. Sci. USA* 105, 10215–10220.
- Burrows, L.L. (2005). Weapons of mass retraction. *Mol. Microbiol.* 57, 878–888.
- CCP4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50, 760–763.
- de la Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* 276, 472–494.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2126–2132.
- Frankel, B.A., Tong, Y., Bentley, M.L., Fitzgerald, M.C., and McCafferty, D.G. (2007). Mutational analysis of active site residues in the *Staphylococcus aureus* transpeptidase SrtA. *Biochemistry* 46, 7269–7278.
- Gianfaldoni, C., Censini, S., Hillergermann, M., Moschioni, M., Facciotti, C., Pansegrau, W., Masignani, V., Covacci, A., Rappuoli, R., Barocchi, M.A., and Ruggiero, P. (2007). *Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge. *Infect. Immun.* 75, 1059–1062.
- Hava, D.L., and Camilli, A. (2002). Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.* 45, 1389–1405.
- Heinig, M., and Frishman, D. (2004). STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* 32, W500–W502.
- Hilleringmann, M., Giusti, F., Baudner, B.C., Masignani, V., Covacci, A., Rappuoli, R., Barocchi, M.A., and Ferlenghi, I. (2008). Pneumococcal pili are composed of protofilaments exposing adhesive clusters of RrgA. *PLoS Pathog.* 4, e1000026.
- Ilangovan, U., Ton-That, H., Iwajara, J., Schneewind, O., and Clubb, R.T. (2001). Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. USA* 98, 6056–6061.
- Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* 26, 795–800.
- Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637.
- Kang, H.J., Coulibaly, F., Clow, F., Proft, T., and Baker, E.N. (2007). Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure. *Science* 318, 1625–1628.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereo chemical quality of protein structures. *J. Appl. Cryst.* 26, 283–291.
- LeMieux, J., Hava, D.L., Basset, A., and Camilli, A. (2006). RrgA and RrgB are components of a multisubunit pilus encoded by the *Streptococcus pneumoniae* rlrA pathogenicity islet. *Infect. Immun.* 74, 2453–2456.
- LeMieux, J., Woody, S., and Camilli, A. (2008). The roles of the sortases of *Streptococcus pneumoniae* in assembly of the RlrA pilus. *J. Bacteriol.* 190, 6002–6013.
- Levine, O.S., O'Brien, K.L., Knoll, M., Adegbola, R.A., Black, S., Cherian, T., Dagan, R., Goldblatt, D., Grange, A., Greenwood, B.M., et al. (2006). Pneumococcal vaccination in developing countries. *Lancet* 367, 1880–1882.
- Maione, D., Margarit, I., Rinaudo, C.D., Masignani, V., Mora, M., Scarselli, M., Tettelin, H., Brettoni, C., Iacobini, E.T., Rosini, R., et al. (2005). Identification of a universal group B *Streptococcus* vaccine by multiple genome screen. *Science* 309, 148–150.
- Mandlik, A., Das, A., and Ton-That, H. (2008). The molecular switch that activates the cell wall anchoring step of pilus assembly in gram-positive bacteria. *Proc. Natl. Acad. Sci. USA* 105, 14147–14152.
- Manetti, A.G.O., Zingaretti, C., Falugi, F., Capo, S., Bombaci, M., Bagnoli, F., Gambellini, G., Bensi, G., Mora, M., Edwards, E.A., et al. (2007). *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. *Mol. Microbiol.* 64, 968–983.
- Marraffini, L.A., and Schneewind, O. (2006). Targeting proteins to the cell wall of sporulating *Bacillus anthracis*. *Mol. Microbiol.* 62, 1402–1417.
- Marraffini, L.A., DeDent, A.C., and Schneewind, O. (2006). Sortases and the art of anchoring proteins to the envelopes of Gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* 70, 192–221.
- Mattick, J.S. (2002). Type IV pili and twitching motility. *Annu. Rev. Microbiol.* 56, 289–314.
- Mazmanian, S.K., Skaar, E.P., Gaspar, A.H., Humayun, M., Gornicki, P., Jelenska, J., Joachimiak, A., Missiakas, D.M., and Schneewind, O. (2003). Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* 299, 906–909.
- Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A.G.O., Maggi, T., Taddei, A.R., Grandi, G., and Telford, J.L. (2005). Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. *Proc. Natl. Acad. Sci. USA* 102, 15641–15646.
- Murshudov, G., Vagin, A., and Dodson, E. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53, 240–255.
- Naik, M.T., Suree, N., Ilangovan, U., Liew, C.K., Thieu, W., Campbell, D.O., Clemens, J.J., Jung, M.E., and Clubb, R.T. (2006). *Staphylococcus aureus* sortase A transpeptidase. Calcium promotes sorting signal binding by altering the mobility and structure of an active site loop. *J. Biol. Chem.* 281, 1817–1826.
- Nallapareddy, S.R., Singh, K.V., Sillanpaa, J., Garsin, D.A., Hook, M., Erlandsen, S.L., and Murray, B.E. (2006). Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. *J. Clin. Invest.* 116, 2799–2807.

- Painter, J., and Merritt, E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D* 62, 439–450.
- Perrakis, A., Morris, R.M., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* 6, 458–463.
- Rosini, R., Rinaudo, C.D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzi, C., Brettoni, C., et al. (2006). Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. *Mol. Microbiol.* 61, 126–141.
- Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815.
- Sauer, F.G., Pinkner, J.S., Waksman, G., and Hultgren, S.J. (2002). Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell* 111, 543–551.
- Schneewind, O., Model, P., and Fischetti, V.A. (1992). Sorting of protein A to the staphylococcal cell wall. *Cell* 70, 267–281.
- Scott, J.R., and Zahner, D. (2006). Pili with strong attachments: Gram-positive bacteria do it differently. *Mol. Microbiol.* 62, 320–330.
- Spratt, B.G., and Greenwood, B.M. (2000). Prevention of pneumococcal disease by vaccination: does serotype replacement matter? *Lancet* 356, 1210–1211.
- Storoni, L., McCoy, A., and Read, R. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* 57, 1373–1382.
- Telford, J.L., Barocchi, M.A., Margarit, I., Rappuoli, R., and Grandi, G. (2006). Pili in Gram-positive pathogens. *Nat. Rev. Microbiol.* 4, 509–519.
- Ton-That, H., and Schneewind, O. (2003). Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol. Microbiol.* 50, 1429–1438.
- Ton-That, H., Mazmanian, S.K., Alksne, L., and Schneewind, O. (2002). Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. *J. Biol. Chem.* 277, 7447–7452.
- Ton-That, H., Maraffini, L.A., and Schneewind, O. (2004). Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol. Microbiol.* 53, 251–261.
- Vetsch, M., Puorger, C., Spirig, T., Grauschopf, U., Weber-Ban, E.U., and Glockshuber, R. (2004). Pilus chaperones represent a new type of protein-folding catalyst. *Nature* 431, 329–332.
- Walsh, C. (2003). Where will new antibiotics come from? *Nat. Rev. Microbiol.* 1, 65–70.
- Weerkamp, A.H., Handley, P., Baars, A., and Slot, J.W. (1986). Negative staining and immunoelectron microscopy of adhesion-deficient mutants of *Streptococcus salivarius* reveal that the adhesive protein antigens are separate classes of cell surface fibril. *J. Bacteriol.* 165, 746–755.
- Yao, M., Zhou, Y., and Tanaka, I. (2006). LAFIRE: software for automating the refinement process of protein structure analysis. *Acta Crystallogr. Sect. D* 62, 189–196.
- Zhang, R., Wu, R., Joachimiak, G., Mazmanian, S.K., Missiakas, D.M., Gornicki, P., Schneewind, O., and Joachimiak, A. (2004). Structures of sortase B from *Staphylococcus aureus* and *Bacillus anthracis* reveal catalytic amino acid triad in the active site. *Structure* 12, 1147–1156.
- Zong, Y., Bice, T.W., Ton-That, H., Schneewind, O., and Narayana, S.V.L. (2004a). Crystal structure of *Staphylococcus aureus* sortase A and its substrate complex. *J. Biol. Chem.* 279, 31383–31389.
- Zong, Y., Mazmanian, S.K., Schneewind, O., and Narayana, S.V.L. (2004b). The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the *Staphylococcus aureus* cell wall. *Structure* 12, 105–112.

Sortase activity is controlled by a flexible lid in the pilus biogenesis mechanism of Gram-positive pathogens

Journal:	<i>Biochemistry</i>
Manuscript ID:	draft
Manuscript Type:	Article
Date Submitted by the Author:	
Complete List of Authors:	Manzano, C; Institut de Biologie Structurale Di Guilmi, Anne Marie; Institut de Biologie Structurale Dessen, Andrea; Institut de Biologie Structurale



Sortase activity is controlled by a flexible lid in the pilus
biogenesis mechanism of Gram-positive pathogens

C. Manzano, A M. Di Guilmi, and A. Dessen*

Institut de Biologie Structurale *Jean-Pierre Ebel*, UMR 5075

(CEA, CNRS, UJF); 41 rue Jules Horowitz, F-38027 Grenoble, France

* to whom correspondence should be addressed

email: dessen@ibs.fr

phone: (33)4-38-78-95-90

fax: (33)4-38-78-54-94

ABBREVIATIONS

TSA, thermal shift assay; LC-MS, liquid chromatography/mass spectrometry

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Pili are surface-linked virulence factors that play key roles in infection establishment in a variety of pathogenic species. In Gram-positive pathogens, pilus formation requires the action of sortases, dedicated transpeptidases that covalently associate pilus building blocks. In *Streptococcus pneumoniae*, a major human pathogen, all genes required for pilus formation are harbored in a single pathogenicity islet which encodes three structural proteins (RrgA, RrgB, RrgC) and three sortases (SrtC-1, SrtC-2, SrtC-3). RrgB forms the backbone of the streptococcal pilus, to which minor pilins RrgA and RrgC are covalently associated. SrtC-1 is the main sortase involved in polymerization of the RrgB fiber and displays a lid which encapsulates the active site, a feature present in all pilus-related sortases. In this work, we show that catalysis by SrtC-1 proceeds through a catalytic triad constituted of His, Arg, and Cys, and that lid instability affects protein fold and catalysis. In addition, we show by thermal shift analysis that lid flexibility can be stabilized by the addition of substrate-like peptides, a feature shared by other periplasmic transpeptidases. We also report the characterization of a trapped acyl-enzyme intermediate formed between SrtC-1 and RrgB. The presence of lid-encapsulated sortases in the pilus biogenesis systems in many Gram-positive pathogens points to a common mechanism of substrate recognition and catalysis that should be taken into consideration in the development of sortase inhibitors.

INTRODUCTION

Pili are extracellular flexible fibers that are directly associated to the bacterial cell wall and play key roles in initiation of bacterial infection. These important virulence factors are essential for adhesion to target host cells, motility, and DNA transformation, and have been shown to considerably enhance bacterial ability to attach to host tissue and initiate colonization^{1; 2}. In contrast to pili located on the surface of Gram-negative bacteria, which are formed by the non-covalent association of subunits^{3; 4}, Gram-positive pili have been shown to be formed by covalently associated building blocks. Pathogens such as *Corynebacterium diphtheriae*, *Streptococcus pyogenes*, and *Streptococcus pneumoniae* carry all of the genes required for generation of the entire pilus-forming machinery within a single pathogenicity islet^{5; 6; 7; 8}. Thirty percent of pathogenic strains of *S. pneumoniae*, the causative agent of pneumonia, septicemia and meningitis, carry the *rlrA* pathogenicity region, which codes for a transcriptional regulator (RlrA), three structural proteins (RrgA, RrgB and RrgC) and three sortases (SrtC-1, SrtC-2 and SrtC-3). Electron microscopy and genetic studies of infectious pneumococcal strains have shown the pilus backbone fiber to be formed by RrgB subunits, while RrgA and RrgC are minor pilins that are directly associated to the polymerized RrgB fiber^{5; 8; 9}. Both RrgA and RrgC play important roles in infectivity and target cell recognition^{9; 10}, while SrtC-1, SrtC-2 and SrtC-3 catalyze the intramolecular association of Rrg molecules^{11; 12; 13}.

Sortases are cysteine transpeptidases that participate in a variety of processes, including surface attachment of virulence factors, sporulation, and iron acquisition^{14; 15}. SrtA and SrtB from *Staphylococcus aureus* represent the best-studied sortases to

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

date. These enzymes recognize an LPXTG-like motif within the sequence of secreted virulence factors, catalyze the nucleophilic attack of the Thr-Gly bond, and finally associate the target proteins carrying the LPXT- sequence directly to the peptidoglycan stem peptide^{16; 17; 18}. The structures of SrtA and SrtB reveal a beta-barrel fold displaying a triad of catalytic residues (Cys, His, Arg) on exposed extremities of three adjacent strands^{19; 20}; the identification of Cys as the catalytic residue responsible for nucleophilic attack was shown not only by mutagenesis studies, in which introduction of an Ala at this site abolishes catalysis, but also by crystallization of covalent complexes between sortases and ligands which are covalently associated to the Cys residue^{21; 22; 23}. In addition, the structure of SrtA in complex with a LPXTG-like peptidyl-sulfhydryl compound reveals a covalent interaction between the peptide and the catalytic Cys²⁴. Notably, while mutagenesis of His or Arg side chains dramatically affects catalytic efficiency in SrtA, the precise role played by these residues is still a matter of controversy^{17; 25; 26; 27}.

Within the pneumococal pilus system, RrgA, RrgB, and RrgC all carry LPXTG-like motifs on their C-termini, suggesting that these regions are targeted by sortases^{5; 28; 29}. Although the equal number of structural and sortase-encoding genes in pneumococci originally implied that each sortase could be exclusively dedicated to one Rrg molecule⁵, recent work has shown that pilus sortases have interchangeable specificities to a certain extent^{11; 30}. Nevertheless, SrtC-1 was shown to be the main sortase involved in polymerization of the RrgB fiber³⁰ while the other two sortases also recognize the coadjuvant adhesin molecules^{5; 11; 12}. The recent structures of two pilus-related sortases (SrtC-1 and SrtC-3 from *S. pneumoniae*) revealed that these enzymes, like their non-pilus-associated counterparts, also fold into beta-barrels and

1
2
3 carry Cys, His, and Arg side chains within the active site cleft. Cys193 was shown to
4
5 be essential for RrgB fiber formation by SrtC-1; the wild type enzyme could generate
6
7 elongated RrgB fibers *in vitro*, while a Cys193Ala mutant showed no activity³⁰.
8
9 Additionally, a noteworthy feature of both pilus-related sortases involves the
10
11 presence of a lid that not only blocks active site access but also carries two key
12
13 residues, which are invariably an Asp and a hydrophobic amino acid, that make
14
15 interactions within the catalytic cleft itself, serving as 'anchors'. Although sequences
16
17 corresponding to lid regions can be identified in all pilus-related sortases
18
19 characterized to date³⁰, the function of this feature, as well as the roles played by the
20
21 two other residues in catalysis, have remained obscure.
22
23
24
25
26
27
28

29 In this work, we investigate the functions of both the lid and catalytic triad
30
31 regions in the pilus fiber-formation activity of SrtC-1. We employ thermal shift assay
32
33 techniques to show that the lid region of SrtC-1 plays roles both in stabilizing the
34
35 enzyme and allowing substrate access to the active site. In addition, we show that
36
37 amino acids His131 and Arg202 are essential for RrgB fiber formation, and that
38
39 Arg202, in particular, also plays a structural stabilization role through interaction
40
41 with lid anchors. Lastly, we describe the 'trapping' of a covalent complex between
42
43 RrgB and SrtC-1, which can be observed by SDS gel and LC-MS/MS spectrometry
44
45 analysis. The absolute requirement for sortases in the pilus-formation processes of all
46
47 Gram-positive pathogens, coupled to the presence of lid regions with similar anchor
48
49 residues in these enzymes, suggest that these data will contribute to the
50
51 understanding of the pilus-formation mechanism in a variety of infectious
52
53 organisms.
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

MATERIALS AND METHODS

Generation of mutants SrtC-1-H131D, SrtC-1-R202E, SrtC-1-D58GW60G, and SrtC-1-C193AD58GW60G

The expression plasmid pLIMSrtC-1 encoding wild-type His6-SrtC-1 (or HisSrtC-1Cys193Ala) served as a template for the introduction of single amino acid substitutions by PCR using the QuickChange XL Site Directed Mutagenesis (Stratagene). All mutant constructs were subsequently sequenced; the results showed that only expected mutations were introduced during PCR.

Expression and purification of recombinant proteins

Protein expression in *E. coli* BL21 (DE3)-RIL was similar in all cases and was induced in Terrific broth with 1 mM IPTG at 37 °C for 3 h. Cells were harvested by centrifugation and lysed by sonication in buffer A (50 mM Tris-HCl, pH 8.0, 0.2 M NaCl, 20 mM imidazole). Supernatants were cleared by centrifugation and applied to a HisTrap column (Amersham Biosciences), pre-equilibrated in buffer A. Proteins were eluted with a imidazole gradient and were subsequently cleaved by Tev protease overnight at 4°C. Recombinant proteins were then submitted to further purification on a gel filtration column in 50 mM HEPES pH 7.0, 0.15 M NaCl. All mutant proteins were verified by mass spectrometry analysis.

***In vitro* RrgB fiber formation**

Recombinant purified RrgB was incubated with SrtC-1, SrtC-1-H131D or SrtC-1-R202E at a 1:4 (w/w) ratio in a total volume of 15 μ l and subsequently incubated overnight at 37°C. Each aliquot was mixed with XT Sample Buffer and XT Reducing Agent (Bio-Rad), boiled for 10 min and loaded onto 4-12% Criterion XT Precast gels (Bio-Rad). Gels were subsequently electro-transferred into Trans-Blot Transfer Medium (Biorad). Membranes were blocked with 5% nonfat milk in PBS with 0.3% Tween (PBS-T) for 1h at room temperature. Membranes were left in primary antiserum (anti-RrgB or anti-SrtC-1 polyclonal mouse antibodies diluted to 1:5,000) during 1.5 hours at RT. After 3 successive washes in PBS-T, the membranes were incubated with secondary anti-mouse HRP-conjugate antibodies (Sigma, diluted 1:10,000) overnight at 4°C. Proteins were detected directly on the membranes with SIGMAFAST™ DAB Tablets (Sigma).

Fluorescence-based protein thermal stability assays

Assays were conducted in an IQ5 96-well format real-time PCR instrument (Bio-Rad) in the presence of a Sypro Orange (Molecular Probes) probe. The total volume was 25 μ L. Samples were heat-denatured from 20°C to 100°C at a rate of 1°C per minute for SrtC-1-H131D and SrtC-1-R202E mutants and from 20°C to 80°C at a rate of 0.5°C per minute for SrtC-1-D8GW60G and SrtC-1-C193AD58GW60G mutants. At each step, excitation was performed at 470 nm, while emission of Sypro Orange fluorescence was monitored at 570 nm. Plotting of the fluorescence vs temperature curves, followed by the calculation of the first derivative at each point allowed the

1
2
3 identification of each inflection point; the minima were referred to as the melting
4 temperatures (T_m). The SrtC-1-C193AD58GW60G variant was employed for
5 experiments with the IPQTG and YPRTG peptides (PSL GmbH), in order to avoid
6 the potential catalytic reaction of the substrates with the active site Cys193. Prior to
7 experimentation with peptides, SrtC-1-C193AD58GW60G and SrtC-1-D58GW60G
8 variants were shown to display identical purification profiles and TSA T_m values (not
9 shown). Peptides were added to the experiment in a final molar ratio of 1:10 and 1:20
10 (protein:peptide). Each result is representative of three independent experiments.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

26 Crystallization, data collection and structure solution

27
28
29
30 SrtC-1-H131D crystals were obtained in hanging drops in 0.1 M Tris pH 8.5,
31 0.2M $MgCl_2$ hexahydrate, 30% PEG4000 at room temperature and were cryoprotected
32 by gradual incubation in increasing amounts of glycerol to a final concentration of
33 20%. Diffraction data were collected at the European Synchrotron Radiation Facility
34 (ESRF) beamline ID29 (Grenoble, France). Diffraction images were indexed and
35 scaled with XDS³¹. The structure was solved by molecular replacement using the
36 program PHASER³² employing the structure of wild type SrtC-1 as a search model.
37 Model building and refinement steps were performed with Coot and REFMAC 5.4³³;
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Mass Spectrometry analysis

1
2
3 The band identified in Fig. 4a with a star was excised from the gel and subsequently
4
5 washed with 50 mM acetonitrile/ammonium bicarbonate 50/50 (v/v) for 30 min,
6
7
8 before dehydration with acetonitrile. Subsequently, the sample was treated with
9
10 7% H₂O₂. After drying was complete, the band was rehydrated in 15 µL digestion
11
12 buffer (50 mM ammonium bicarbonate pH 8.1) containing 150 ng of trypsin and
13
14 incubated at 4°C for 15 mins. 30µL of digestion buffer were then added and the
15
16 digestion reaction was carried out at 37°C overnight. Peptides were extracted from
17
18 the gel by diffusion for 15 min with agitation, followed by three sequential 5 min
19
20 sonication steps in 50% acetonitrile, 5% formic acid, and 100% acetonitrile. Digestion
21
22 and extraction solutions were pooled and dried under vacuum. Peptide mixtures
23
24 were re-dissolved in 25µl of water/acetonitrile 95/5 (v/v) containing 0.2%
25
26 Trifluoroacetic acid prior to LCMS/MS analysis. All experiments were
27
28 performed in a 96-well system coordinated by an EVO15 (Tecan) robot (EdyP,
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60 Grenoble).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

RESULTS

RrgB fiber formation requires a catalytic triad

SrtC-1 is the main sortase involved in formation of the RrgB pilus fiber, and mutation of its catalytic cysteine (Cys193) into Ala abolishes fiber-forming activity¹¹; ³⁰. The high resolution structure of SrtC-1 from infectious pneumococcal strain TIGR4 revealed that, in addition to Cys193, the catalytic site also harbors His131 and Arg202; although the functions of the two latter residues were unknown, their three-dimensional arrangement is comparable to what is observed in active sites of non-pilus-related sortases, which also carry Cys, His, and Arg residues^{19; 22; 24; 30}. In order to explore the functions of His131 and Arg202 in SrtC-1's fiber-forming activity, we mutated these residues into Asp and Glu, respectively, and searched for the capacity of SrtC-1-H131D and SrtC-1-R202E variants to generate a pattern of high molecular weight bands on SDS-PAGE upon incubation with RrgB at 37°C (as described in³⁰). The Western blot on Fig. 1a, which was developed with anti-RrgB antibodies, confirms that wild type SrtC-1 is able to catalyze the formation of a pattern of high molecular weight RrgB bands (which correspond to the pilus backbone fibers as shown by microscopy by Manzano and coworkers³⁰). In contrast, neither SrtC-1-H131D (lane 3) nor SrtC-1-R202E (lane 4) were able to catalyze fiber formation. These results suggest that, as in the case of non-pilus-related sortases, SrtC-1 employs a catalytic triad composed of Cys193, His131, and Arg202 for RrgB fiber generation.

In order to address the issue of active site stability within SrtC-1 mutant structures, we submitted both proteins to crystallization trials. SrtC-1-R202E crystals were of poor quality and diffracted X-rays only to low resolution (8 Å). However, SrtC-1-H131D crystals diffracted X-rays to high resolution (1.3 Å) at the ESRF synchrotron in Grenoble and we were able to solve its structure by employing wild type SrtC-1 (PDB code 2WIJ) as a search model in a molecular replacement experiment (see methods). Data collection and refinement statistics are shown in Table I. SrtC-1-H131D displays a highly similar fold to wild type SrtC-1 (rms deviation of 0.2 Å over 196 Cα atoms), revealing that the mutation in the catalytic cleft does not affect the general fold of the protein. Analysis of the SrtC-1-H131D active site (Fig. 1c, where lid residues are shown in green) reveals only slight differences from the wild-type SrtC-1 molecule (Fig. 1b). The side chain of Cys193 is present in two conformations, as is the case for wild type SrtC-1 (only one is shown for clarity). In addition, the side chain of the mutant residue, Asp131, stabilizes the Nε group of Trp60, a lid anchor residue, through a close hydrogen bond (2.7 Å); in the wild type structure, His131 does not make any interactions within the active site, and is located 5.3 Å away from the nucleophilic cysteine. Thus, although the mutation of His131 into Asp does not affect the overall fold of the protein, Asp131 is not able to replace the catalytic function of His131, suggesting that the positive charge on His131 is crucial for SrtC-1's fiber-forming activity.

We further investigated the stability of wild type and mutant SrtC-1 forms through thermal shift assays. Proteins were gradually heated in a IQ5 real time PCR apparatus and thermal unfolding curves were monitored through the detection of

changes in fluorescence of a Sypro Orange probe. For data analysis, the melting temperature (T_m) for each sample was determined as the minimum peak in the first derivative plot of the melting curve representing the raw fluorescence data. Results are shown in Fig. 2a. This analysis reveals that the T_m values for both wild type SrtC-1 and SrtC-1-H131D are in the range of 46°C, while that of the SrtC-1-R202E mutant is of 39°C. This suggests that mutation of Arg202 engenders structural destabilization of SrtC-1, which could explain the lack of success in obtaining well-diffracting crystals of this mutant. It is of note that in the active site of the wild type SrtC-1 molecule, lid anchor residue Asp58 interacts directly with both NH₂ and N ϵ atoms of Arg202 (Fig. 1b). The results with the SrtC-1-R202E mutant thus suggest that, at least in the absence of substrate, the Asp58-Arg202 interaction is essential for lid and overall protein stability. Thus, the lack of fiber-forming activity of SrtC-1-R202E could be linked not only to the absence of the Arg side chain for catalytic function (discussed below), but also to a decrease in structural stability.

The sortase lid is key for activity and stability

In all pilus-related sortases sequenced to date, lid anchor residues invariably include an Asp at a position which is homologous to that occupied by Asp58 in SrtC-1, and a large hydrophobic side chain (Trp, Tyr, or Phe) at the site occupied by Trp60 (Fig. 2b). The structure of SrtC-1 shows that lid anchor residues not only interact directly with active site residues, as seen above, but the lid itself blocks an elongated cleft which could be the LPXTG substrate binding site³⁰. This hypothesis was put forth through comparison of the structures of SrtC-1/SrtC-3 and staphylococcal SrtA,

1
2
3 which binds an LPXTG peptidic substrate in a comparable cleft that harbors the
4
5 active site residues on one end^{23, 30}, and has also been suggested by the recent
6
7 structure of SrtA in complex with a covalent substrate analog²⁴. This could imply
8
9 that one of the functions of the lid is to regulate substrate access into the cleft. In
10
11 order to investigate this potential function of the lid, we constructed a SrtC-1 double
12
13 mutant in which lid anchor residues Asp58 and Trp60 were mutated into glycines.
14
15 Initially, we tested the ability of SrtC-1-D58GW60G to generate a RrgB high-
16
17 molecular-weight ladder on SDS-PAGE. As shown in the Western blot on Fig. 3a, in
18
19 which the same amount of sample mixture was loaded in all lanes, SrtC-1-
20
21 D58GW60G (lane 2) can recognize RrgB and generate an initial set of bands, but
22
23 polymerization is less efficient than for wild type SrtC-1 (lane 3).
24
25
26
27
28
29
30
31
32

33 To explore the effects of the two anchor side chains on the stability of SrtC-1,
34
35 we performed thermal shift assay (TSA) experiments on SrtC-1-D58GW60G. The
36
37 results, shown in Fig. 3b, indicate that SrtC-1-D58GW60G displays a T_m which is 8°C
38
39 lower than that of the wild type enzyme (38°C versus 46°C). Thus, the side chains of
40
41 lid anchor residues Asp58 and Trp60 are required for SrtC-1 stability. It is of note
42
43 that the SrtC-1-R202E mutant is also considerably less stable than wild type SrtC-1
44
45 (Fig. 2). This indicates that the Arg202-Asp58 interaction plays an important role in
46
47 SrtC-1 stabilization and suggests that mutants that cannot generate this interaction
48
49 have a lid region that tends to be in an 'open', or flexible state. If so, these mutant
50
51 forms could potentially be stabilized by the addition of substrate into the active site
52
53 region.
54
55
56
57
58
59
60

1
2
3 In order to test this possibility, we employed the TSA technology to test if
4
5 SrtC-1-D58GW60G could be stabilized by a peptide corresponding to the RrgB
6
7 LPXTG-like sequence (IPQTG). Addition of increasing amounts of peptide were
8
9 accompanied by an increase in the T_m of the mutant protein (Fig. 3b), suggesting that
10
11 the 5-residue substrate peptide can penetrate the active site of the 'open lid' SrtC-1
12
13 form and stabilize its structure. These data thus strongly point to a double role for
14
15 the lid region in pilus-related sortases: (1) covering of the active site in the absence of
16
17 substrate and ensuring protein stability (2) maintenance of Arg202 optimally pointed
18
19 towards the active site cysteine (through the Asp58-Arg202 interaction), enabling it
20
21 to efficiently participate in catalysis. The presence of a Phe and Asp residues as
22
23 anchor points in the lid region of SrtC-3³⁰, also expressed in the *rlrA* operon, suggests
24
25 that lids in pilus-related sortases could play similar roles.
26
27
28
29
30
31
32
33
34

35 Recently, Falker and co-workers¹¹ showed that SrtC-1 (named SrtB in their
36
37 manuscript), in addition to acting as a pilus polymerase, could also recognize minor
38
39 pilins. We thus set out to verify this point by incubating SrtC-1-D58GW60G with
40
41 increasing amounts of YPRTG, the LPXTG-like peptide from RrgA, a minor pilin,
42
43 and performing analysis through TSA, as described for the IPQTG peptide. Results
44
45 show that, much as in the case of IPQTG, SrtC-1-D58GW60G can also recognize and
46
47 is also stabilized by increasing amounts of YPRTG (Fig. 3c). The T_m is thus shifted
48
49 from 38°C (mutant enzyme with no peptide) to 43°C (highest amount of peptide),
50
51 indicating that the peptide from a minor pilin is as successful as IPQTG in rendering
52
53 SrtC-1-D58GW60G more stable. Thus, the presence of proline, threonine, and glycine
54
55 within the LPXTG-like sequence plays a key role in recognition of the motif within
56
57
58
59
60

the sortase active site. Interestingly, Suree and coworkers²⁴ have recently proposed that the lid anchor residues in pilus-related sortases (Asp58 and Trp60 in SrtC-1) mimic the PX portion of the peptide motif and thus occlude the LPXTG binding site. Our results are in agreement with this proposition, since the mutant that lacks both anchor residues can recognize/be stabilized by two slightly different LPXTG-like sequences. These data also suggest that pilus-related sortases do not show strict specificity for one specific LPXTG-like sequence, as proposed before^{11; 30}. Preferences for the individual Rrg substrates could potentially be dictated by the section of the lid region that faces the outside of the molecule (and not the active site)³⁰, and/or the region of the cleft which recognizes the donor substrate²⁴.

Formation *in vitro* of a covalent complex between SrtC-1 and RrgB

Upon incubation of SrtC-1 and RrgB and analysis on SDS-PAGE, in addition to the RrgB ladder identified with anti-RrgB antibodies (Fig. 1, lane 2), development of the gel with an antibody raised against SrtC-1 revealed three bands (Fig. 4a). The most prominent one represents the sortase monomer, whilst an intermediary band, identifiable both in the presence or absence of RrgB, represents a minor form of a sortase dimer. The other additional band migrates above the 85 kDa marker and is indicated with star in Fig. 4a; interestingly, a band migrating at approximately the same point can also be identified in a gel developed with anti-RrgB serum (Fig. 1 and³⁰), indicating that it could represent a complex between RrgB and SrtC-1. In order to investigate the nature of this band, we submitted the entire mix to mass spectrometry, and the results are shown in Fig. 4b. In addition to the expected

masses of SrtC-1 and RrgB, we identified a form whose molecular mass (approx. 88.3 kDa) could correspond to that of a covalent complex between SrtC-1 and RrgB. In order to confirm this finding, we cleaved the band from the gel, trypsinized the sample, and performed liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). The peptides that correspond to this experiment are shown in Fig. 4c, and all correspond to either RrgB or SrtC-1, indicating that this additional band corresponds to a covalent complex between RrgB and SrtC-1. Notably, the same band could be identified in experiments performed with SrtC-1-D58GW60G, which was shown (Fig. 3a) to display some RrgB-polymerizing capabilities, but no evidence of the complex could be detected with mutants SrtC-1-C193A, SrtC-1-H131D or SrtC-1-R202E (data not shown). These results indicate that the RrgB polymerization reaction proceeds through an acyl-complex intermediate whose formation is dependent on an intact catalytic triad, and which can be trapped *in vitro*.

DISCUSSION

Pilus biosynthesis in Gram-positive bacteria requires the action of dedicated sortases in order to guarantee the covalent association of its building blocks³⁵. The pilus biosynthesis *rlrA* pathogenicity islet of *S. pneumoniae* encodes three sortase enzymes (SrtC-1, SrtC-2, SrtC-3) that do not display strict substrate specificities towards Rrg molecules, although SrtC-1 has been shown to be the key protein in formation of the RrgB fiber backbone^{11;30}. In this work, we have shown that, as is the case for non-pilus-related sortases, a catalytic triad composed of Cys, Arg, and His residues are essential for the catalytic activity of SrtC-1. Although the side chain of His131 is located more than 5 Å away from the nucleophilic Cys193 and its mutation into Asp does not affect protein stability, this modification completely blocks SrtC-1's fiber formation ability. Interestingly, mutation of Arg202 into Glu, in addition to generating a catalytically inactive SrtC-1, also affects enzyme stability, most probably due to the loss of two hydrogen bonds with an anchor lid residue (Asp58; see below).

We were also able to identify a 'trapped' RrgB-SrtC-1 covalent intermediate in the reaction catalyzed by wild type SrtC-1 by Western blotting and LC-MS/MS. This suggests that, although the RrgB fiber can be formed *in vitro*, the SrtC-1 RrgB polymerization reaction in these conditions is slow and/or inefficient, allowing the intermediate to be identified. The covalent intermediate was also identified in experiments in which the SrtC-1-D58GW60G mutant was incubated with RrgB; this indicates that mutation of the lid anchor residues does not block either substrate recognition or acyl-enzyme formation, but affects reaction processivity. Notably, no

bands for the covalent intermediate could be identified in SrtC-1 mutants in which the His131 or Arg202 side chains were mutated. The absence of fiber-formation activity and the inability of Arg and His active site mutants to form the acyl-enzyme complex suggest that both residues participate in a catalytic step prior to acyl-enzyme formation. In SrtA, the catalytic His residue has been hypothesized as being responsible for the protonation of the substrate leaving group, while the Arg side chain could stabilize the negative charge accumulation in the transition state^{21; 25; 36} or participates in positioning of the substrate by direct interaction with backbone residues²⁴. Thus, acyl-enzyme trapping can be enhanced by mutation of either of the two auxiliary members of the catalytic triad, His or Arg. It is also of note that recently Ton-That and co-workers identified pilus polymerization acyl-enzyme intermediates on the surface of *Corynebacterium diphtheriae* mutants³⁷, underlining that the formation of a covalent complex in the pilus biogenesis process of Gram-positive species is also identifiable directly on the bacterial surface.

Sortases involved in pilus formation in a variety of bacterial species carry lid regions with two anchor residues (Fig. 2b). In SrtC-1, these key residues (Asp58 and Trp60) anchor the lid in a seemingly closed conformation. Since, in the structure of SrtC-3, Trp60 is replaced by a Phe³⁰, it seems likely that the hydrophobic nature of this residue plays a role in lid closure. Asp58, however (which is invariably an Asp in all sequences of pilus-related sortases known to date; examples in Fig. 2b), plays a key role in stabilizing Arg202 (thus, a member of the catalytic triad) through two hydrogen bonds, with the side chain facing the Cys nucleophile directly. Absence of the two lid anchor side chains not only starkly diminishes SrtC-1's capacity to

polymerize RrgB fibers, but also causes protein destabilization. Interestingly, protein stability can be partly recovered through the addition of the LPXTG-like peptides, which have been suggested to lodge within the elongated cleft upon lid opening. These results support the proposition that lid movement is required prior to/concomitant with substrate accommodation within the binding site³⁰. The observation that large amounts of the LPXTG-like peptides cannot totally recover SrtC-1 stability suggests that other features of the full-length RrgB substrate also play a role in SrtC-1 recognition and stabilization.

The presence of a flexible lid that covers the active site in pilus-related sortases, studied here, is also detected in other transpeptidases. Suree and coworkers²⁴ have identified that the access of substrate to the active site of SrtA from *S. aureus*, an enzyme involved in the covalent association of virulence factors to the peptidoglycan, is also regulated by a flexible loop. Although this region corresponds to a loop which is in a different position from the lid described in this work for sortases involved in pilus formation (Fig. 2b), it points to the possibility that sortases control access of substrate to their active site through the movement of surrounding flexible regions. Interestingly, another class of well-studied transpeptidases, Penicillin-Binding Proteins (PBPs), also carry a malleable lid region in close proximity to the active site (the loop between $\beta 3$ and $\beta 4$;^{38; 39; 40; 41; 42}). The transpeptidase domains of PBPs catalyze the cross-linking of stem peptides of the peptidoglycan, thus recognizing peptidic substrates within the cleft, much like sortases; in addition, they are the targets for β -lactam antibiotics, whose structure mimics the last two residues of the stem peptide⁴³. Thus, it is of interest that in the crystal structure of PBP1b from

1
2
3 *S. pneumoniae*, the $\beta 3/\beta 4$ loop maintains the active site in 'closed' conformation in the
4
5 absence of ligand largely through its obstruction by a polar residue (Asn656)³⁸.
6
7 Notably, opening of the cleft can only be achieved *in vitro* through incubation with
8
9 substrate or a pseudo-substrate which structurally mimics the pentapeptide. This
10
11 incubation, which causes an 'opening' movement of the loop, allows trapping of an
12
13 acyl-enzyme intermediate as well as complex formation with different β -lactam and
14
15 γ -lactam antibiotics, suggesting that substrate/pseudo-substrate recognition and lid
16
17 opening are related events^{44; 45}. Thus, due to the similarity in substrate recognition
18
19 patterns and catalytic functionality, a common mechanism of regulation of active site
20
21 access could be proposed for both sortase and PBP transpeptidases. In the absence of
22
23 substrate (stem peptides for PBPs, LPXTG-like peptides for sortases) the catalytic
24
25 cleft is maintained in closed conformation through specific interactions generated by
26
27 lid residues (Fig. 5). However, once substrate becomes available (i.e., when
28
29 peptidoglycan biosynthesis is required during different moments of the cell cycle
30
31 (PBPs) or when virulence factors subunits must be attached to partner molecules
32
33 (sortases), the lid is opened, potentially through direct interaction with substrates.
34
35 These observations lend support to the idea that conformational changes induced by
36
37 ligand binding should be taken into consideration in transpeptidase-targeted drug
38
39 design efforts.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ACKNOWLEDGEMENTS

The coordinates of SrtC-1-H131D were deposited in the Protein Databank with code xx and are available pre-release from the authors. The authors wish to thank Carlos Contreras-Martel (IBS) for help during the initial phases of the project, the ESRF ID29 beamline staff for help with data collection, Izabel Bérard and Eric Forest from the IBS mass spectroscopy facility (LSMP, IBS) for analyses, and Alexandra Kraut and Jérôme Garin (Lab. d'Etude de la Dynamique des Protéomes, EDyP Grenoble) for access to the LC-MS/MS platform and experiments. This work was supported by EC grant LSHM-CT-2004-512138 (to A.D).

REFERENCES CITED

1. Telford, J. L., Barocchi, M. A., Margarit, I., Rappuoli, R. & Grandi, G. (2006). Pili in Gram-positive pathogens. *Nature Rev. Microbiol.* 4, 509-519.

2. Proft, T. & Baker, E. N. (2009). Pili in Gram-negative and Gram-positive bacteria -- structure, assembly, and their role in disease. *Cell Mol. Life Sci.* 66, 613-635.

3. Sauer, F. G., Pinkner, J. S., Waksman, G. & Hultgren, S. J. (2002). Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell* 111, 543-551.

4. Vetsch, M., Puorger, C., Spirig, T., Grauschopf, U., Weber-Ban, E. U. & Glockshuber, R. (2004). Pilus chaperones represent a new type of protein-folding catalyst. *Nature* 431, 329-332.

5. LeMieux, J., Hava, D. L., Basset, A. & Camilli, A. (2006). RrgA and RrgB are components of a multisubunit pilus encoded by the *Streptococcus pneumoniae* rlrA pathogenicity islet. *Infect. Immun.* 74, 2453-2456.

6. Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A. G. O., Maggi, T., Taddei, A. R., Grandi, G. & Telford, J. L. (2005). Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. *Proc. Natl. Acad. Sci. USA* 102, 15641-15646.

7. Ton-That, H., Marraffini, L. A. & Schneewind, O. (2004). Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol. Microbiol.* 53, 251-261.

8. Barocchi, M. A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernebro, J., Moschioni, M., Masignani, V., Hultenby, K., Taddei, A. R., Beiter, K., Wartha, F., von Euler, A., Covacci, A., Holden, D. W., Normark, S., Rappuoli, R. & Henriques-Normark, B. (2006). A pneumococcal pilus influences virulence and host inflammatory responses. *Proc. Natl. Acad. Sci. U S A* 103, 2857-2862.

9. Hillerlingmann, M., Giusti, F., Baudner, B. C., Masignani, V., Covacci, A., Rappuoli, R., Barocchi, M. A. & Ferlenghi, I. (2008). Pneumococcal pili are composed of protofilaments exposing adhesive clusters of RrgA. *PLoS Pathogens* 4, e1000026.

10. Nelson, A. L., Ries, J., Bagnoli, F., Dahlberg, S., Falkner, S., Rounioja, S., Tschop, J., Morfeldt, E., Ferlenghi, I., Hillerlingmann, M., Holden, D. W., Rappuoli, R., Normark, S., Barocchi, M. A. & Henriques-Normark, B. (2007). RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol. Microbiol.* 66, 329-340.

11. Fälker, S., Nelson, A. L., Morfeldt, E., Jonas, K., Hultenby, K., Ries, J., Melefors, O., Normark, S. & Henriques-Normark, B. (2008). Sortase-mediated assembly and surface topology of adhesive pneumococcal pili. *Mol. Microbiol.* 70, 597-607.

12. LeMieux, J., Woody, S. & Camilli, A. (2008). The roles of the sortases of *Streptococcus pneumoniae* in assembly of the RlrA pilus. *J. Bacteriol.* e-print.

13. Dramsi, S., Magnet, S., Davison, S. & Arthur, M. (2008). Covalent attachment of proteins to peptidoglycan. *FEMS Microbiol. Rev.* 32, 307-320.

14. Marraffini, L. A., DeDent, A. C. & Schneewind, O. (2006). Sortases and the art of anchoring proteins to the envelopes of Gram-positive bacteria. *Microb. Mol. Biol. Rev.* 70, 192-221.

15. Mazmanian, S. K., Skaar, E. P., Gaspar, A. H., Humayun, M., Gornicki, P., Jelenska, J., Joachimiak, A., Missiakas, D. M. & Schneewind, O. (2003). Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* 299, 906-909.

16. Schneewind, O., Model, P. & Fischetti, V. A. (1992). Sorting of protein A to the staphylococcal cell wall. *Cell* 70, 267-281.
17. Ton-That, H., Mazmanian, S. K., Alksne, L. & Schneewind, O. (2002). Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. *J. Biol. Chem.* 277, 7447-7452.
18. Mandlik, A., Swierczynski, A., Das, A. & Ton-That, H. (2008). Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol.* 16, 33-40.
19. Zhang, R., Wu, R., Joachimiak, G., Mazmanian, S. K., Missiakas, D. M., Gornicki, P., Schneewind, O. & Joachimiak, A. (2004). Structures of sortase B from *Staphylococcus aureus* and *Bacillus anthracis* reveal catalytic amino acid triad in the active site. *Structure* 12, 1147-1156.
20. Ilangovan, U., Ton-That, H., Iwajara, J., Schneewind, O. & Clubb, R. T. (2001). Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. USA* 98, 6056-6061.
21. Maresso, A. W., Wu, R., Kern, J. W., Zhang, R., Janik, D., Missiakas, D. M., Duban, M.-E., Joachimiak, A. & Schneewind, O. (2007). Activation of inhibitors by sortase triggers irreversible modification of the active site. *J. Biol. Chem.* 282, 23129-23139.
22. Zong, Y., Mazmanian, S. K., Schneewind, O. & Narayana, S. V. L. (2004). The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the *Staphylococcus aureus* cell wall. *Structure* 12, 105-112.
23. Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O. & Narayana, S. V. L. (2004). Crystal structure of *Staphylococcus aureus* sortase A and its substrate complex. *J. Biol. Chem.* 279, 31383-31389.
24. Suree, N., Liew, C. K., Villareal, V. A., Thieu, W., Fadeev, E. A., Clemens, J. J., Jung, M. E. & Clubb, R. T. (2009). The structure of the *Staphylococcus aureus* sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized. *J. Biol. Chem.* e-print.
25. Frankel, B. A., Tong, Y., Bentley, M. L., Fitzgerald, M. C. & McCafferty, D. G. (2007). Mutational analysis of active site residues in the *Staphylococcus aureus* transpeptidase SrtA. *Biochemistry* 46, 7269-7278.
26. Marraffini, L. A., Ton-That, H., Zong, Y., Narayana, S. V. L. & Schneewind, O. (2004). Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*: a conserved arginine residue is required for efficient catalysis of sortase A. *J. Biol. Chem.* 279, 37763-37770.
27. Liew, C. K., Smith, B. T., Pilpa, R., Suree, N., Ilangovan, U., Connolly, K. M., Jung, M. E. & Clubb, R. T. (2004). Localization and mutagenesis of the sorting signal binding site on sortase A from *Staphylococcus aureus*. *FEBS Lett.* 571, 221-226.
28. Scott, J. R. & Zahner, D. (2006). Pili with strong attachments: Gram-positive bacteria do it differently. *Mol. Microbiol.* 62, 320-330.
29. Hava, D. L. & Camilli, A. (2002). Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.* 45, 1389-1405.
30. Manzano, C., Contreras-Martel, C., El Mortaji, L., Izore, T., Fenel, D., Vernet, T., Schoehn, G., Di Guilmi, A. M. & Dessen, A. (2008). Sortase-mediated pilus fiber biogenesis in *Streptococcus pneumoniae*. *Structure* 16, 1838-1848.
31. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* 26, 795-800.
32. Storoni, L., McCoy, A. & Read, R. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr. sect. D* 57, 1373-1382.

- 1
2
3 33. Murshudov, G., Vagin, A. & Dodson, E. (1997). Refinement of macromolecular
4 structures by the maximum-likelihood method. *Acta Crystallogr. sect. D* 53, 240-255.
5
6 34. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics.
7 *Acta Crystallogr. D* 60, 2126-2132.
8
9 35. Ton-That, H. & Schneewind, O. (2004). Assembly of pili in Gram-positive bacteria.
10 *Trends Microbiol.* 12, 228-234.
11
12 36. Frankel, B. A., Kruger, R. G., Robinson, D. A., Kelleher, N. L. & McCafferty, D. G.
13 (2005). *Staphylococcus aureus* sortase transpeptidase SrtA: insight into the kinetic
14 mechanism and evidence for a reverse protonation catalytic mechanism. *Biochemistry*
15 44, 11188-11200.
16
17 37. Guttilla, I. K., Gaspar, A. H., Swierczynski, A., Swaminathan, A., Dwivedi, P., Das,
18 A. & Ton-That, H. (2009). Acyl enzyme intermediates in the sortase-catalyzed pilus
19 morphogenesis in Gram-positive bacteria. *J. Bacteriol.* eprint.
20
21 38. Macheboeuf, P., Di Guilmi, A. M., Job, V., Vernet, T., Dideberg, O. & Dessen, A.
22 (2005). Active site restructuring regulates ligand recognition in class A penicillin-
23 binding proteins. *Proc. Natl. Acad. Sci. USA* 102, 577-582.
24
25 39. Sauvage, E., Kerff, F., Fonce, E., Herman, R., Schoot, B., Marquette, J. P., Taburet,
26 Y., Prevost, D., Dumas, J., Leonard, G., Stefanic, P., Coyette, J. & Charlier, P. (2002).
27 The 2.4 Å crystal structure of the penicillin-resistant penicillin-binding protein
28 PBP5fm from *Enterococcus faecium* in complex with benzylpenicillin. *Cell Mol. Life*
29 *Sci.* 59, 1223-1232.
30
31 40. Powell, A. J., Tomberg, J., Deacon, A. M., Nicholas, R. A. & Davies, C. (2008).
32 Crystal structures of penicillin-binding protein 2 from penicillin-susceptible and -
33 resistant strains of *N. gonorrhoeae* reveal an unexpectedly subtle mechanism for
34 antibiotic resistance. *J. Biol. Chem.* e-print.
35
36 41. Contreras-Martel, C., Dahout-Gonzalez, C., Dos Santos Martins, A., Kotnik, M. &
37 Dessen, A. (2009). PBP active site flexibility as the key mechanism for beta-lactam
38 resistance in pneumococci. *J. Mol. Biol.* 387, 899-909.
39
40 42. Contreras-Martel, C., Job, V., Di Guilmi, A. M., Vernet, T., Dideberg, O. & Dessen,
41 A. (2006). Crystal structure of Penicillin-Binding Protein 1a (PBP1a) reveals a
42 mutational hotspot implicated in β -lactam resistance in *Streptococcus pneumoniae*. *J.*
43 *Mol. Biol.* 355, 684-696.
44
45 43. Hölte, J. V. (1998). Growth of the stress-bearing and shape-maintaining murein
46 sacculus of *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* 62, 181-203.
47
48 44. Macheboeuf, P., Lemaire, D., Thaller, N., Dos Santos Martins, A., Luxen, A.,
49 Dideberg, O., Jamin, M. & Dessen, A. (2007). Trapping of an acyl-enzyme
50 intermediate in a penicillin-binding protein (PBP) catalyzed reaction. *J. Mol. Biol.*
51 376, 405-413.
52
53 45. Macheboeuf, P., Fischer, D. S., Brown Jr., T., Zervosen, A., Luxen, A., Joris, B.,
54 Dessen, A. & Schofield, C. J. (2007). Structural and mechanistic basis of penicillin-
55 binding protein inhibition by lactvicins. *Nat. Chem. Biol.* e-print.
56
57
58
59
60

TABLE I

SrtC-1H131D	
Data Collection	
X-ray source	ID29
Detector	ADSC Q315
Space Group	P2 ₁ 2 ₁ 2 ₁
Unit cell parameters (Å)	a= 68.4, b = 70.5, c = 87.3
Wavelength (Å)	0.980880
Resolution (Å)	3.9-1.3 (1.4-1.3)
No. of unique reflections	99255 (14789)
Completeness (%)	95.3 (88.9)
Mosaicity (°)	0.16
R _{sym}	5.4 (36.4)
I/σ(I)	17.0 (4.2)
Refinement	
Resolution (Å)	19.8-1.3
R _{work} (%)	0.16
R _{free} (%)	0.20
No. of protein atoms	3286
No. of solvent atoms	670
No. of glycerol molecules	3
RMS deviation, bond lengths (Å)	0.01
RMS deviation, bond angles(°)	1.41
Mean B-factor (Å ²)	14.99
Ramachandran favoured (%)	98.7

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE LEGENDS

Fig. 1. The active site of SrtC-1 contains a catalytic triad.

(A) Western blot developed with RrgB antiserum. Lanes: (1), monomeric RrgB; (2), RrgB + SrtC-1; (3), RrgB + SrtC-1-H131D; (4), RrgB + SrtC-1-R202E. Molecular mass standards (in kDa) are indicated on the left. (B) active site of wild type SrtC-1 and (C) active site of SrtC-1-H131D. The active site residues are shown as sticks and the lid region is colored green.

Fig.2 . Thermal stability of active site SrtC-1 variants (A) Thermal unfolding of SrtC-1 and two active site mutants was followed in the presence of the Sypro Orange fluorescent probe (left). The first derivative of the thermal denaturation curve allows the determination of the minima reflecting the melting temperature (right). (B) Sequence alignment of the lid regions of sortases involved in pilus formation in different Gram-positive pathogens (six first lines) and association of virulence factors to the peptidoglycan (two last lines). Those involved in pilus biosynthesis harbor the Asp-Pro-Hyd motif, in which the Asp and the Hydrophobic residue serve as anchors within the catalytic cleft.

Fig. 3. Activity and thermal stability studies of SrtC-1 carrying mutations in the lid anchor residues.

(A) Western blot developed with antiserum against RrgB shows that SrtC-1-D58W60G polymerizes RrgB less efficiently. Lanes: (1), monomeric RrgB; (2), RrgB +

SrtC-1-D58GW60G; (3), RrgB + wild type SrtC-1. The values on the left correspond to molecular weight marker sizes in kDa. (B) Thermal unfolding studies of SrtC-1-D58W60GC193A in the presence of IPQTG, which corresponds to RrgB's LPXTG-like motif. The incubation of SrtC-1-D58W60GC193A with the IPQTG peptide at 2 increasing molar ratios (1:10 and 1:20, protein:peptide) suggests a stabilization of the mutant by the peptide. (C) Thermal unfolding studies of SrtC-1-D58W60GC193A in the presence of YPRTG, which corresponds to RrgA's LPXTG-like motif. In both (B) and (C), the right panels represent the first derivative of the denaturation curves shown on the left.

Fig. 4. SrtC-1 and RrgB form a covalent complex

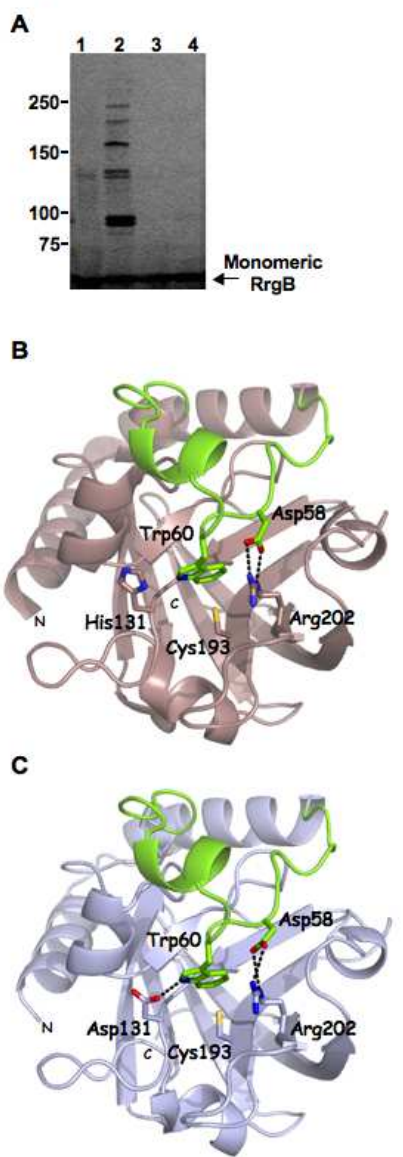
(A) Western blot developed with SrtC-1 antiserum. Lanes: (1), monomeric SrtC-1; (2), RrgB + SrtC-1. Monomeric SrtC-1 migrates as a band with an approximate molecular mass of 23 kDa; a small amount of dimeric SrtC-1, which migrates above its theoretical mass of 47.4 kDa, is also visible. The RrgB/SrtC-1 covalent complex is indicated with a star. (B) MALDI results of the mix between SrtC-1 and RrgB, which identifies not only the single proteins but also a mass of approximately 88 kDa, which corresponds to a covalent complex. On occasion, the SrtC-1 dimer could also be identified, yielding a mass value of 47,573 Da. (C) LC-MS/MS results for the band indicated with a star in (A). The lanes correspond to: *observed*: mass to charge ratio (m/z) of peptide; *Mr (expt)*: experimental molecular mass (Da) of peptide; *Mr (calc)*: calculated molecular mass (Da) of peptide; *Delta*: mass difference (Da) between *Mr (expt)* and *Mr (calc)*; *First*: identification of the first amino acid of the peptide in the full-length protein; *Last*: identification of the last amino acid of the peptide on the

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

full- length protein; *Sequence*: sequence of the identified peptide.

Fig. 5. Access to transpeptidation active sites in PBPs and sortases is regulated by flexible lids. Surface representations of (A) PBP1b from *S. pneumoniae* and (B) SrtC-1, with active site residues represented in red. Both transpeptidases display lid regions (yellow for PBP1b, green for SrtC-1) which are maintained in closed conformation in the absence of ligand. Polar residues Asn656 (A) and Asp58 (B) play key roles in maintaining lids in a closed form.

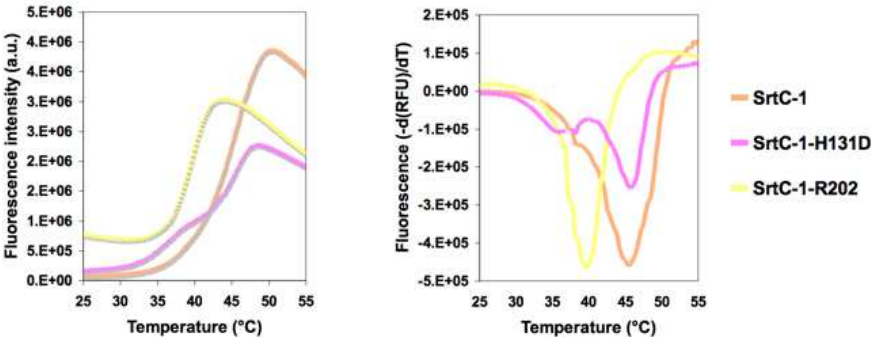
Figure 1



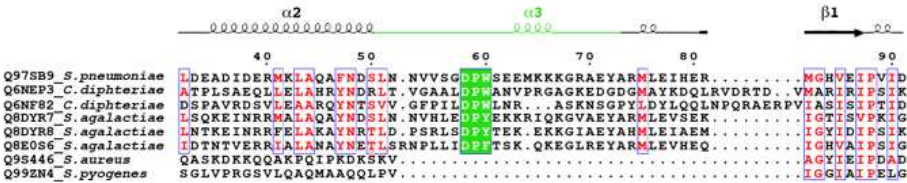
100x286mm (72 x 72 DPI)

Figure 2

A



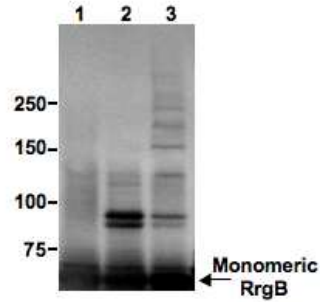
B



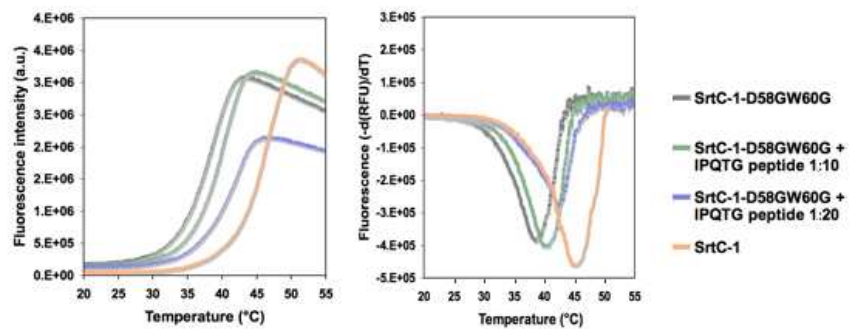
259x220mm (72 x 72 DPI)

Figure 3

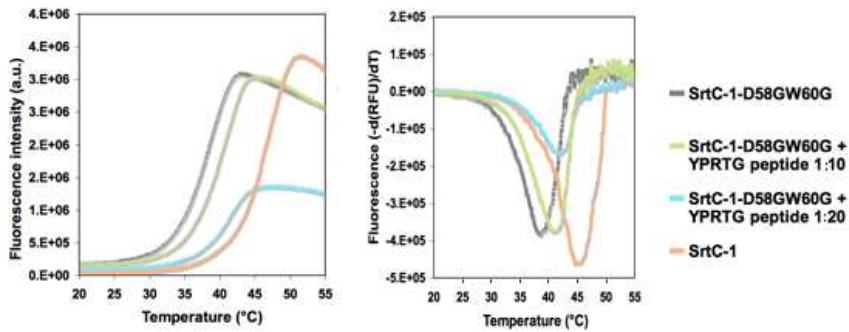
A



B

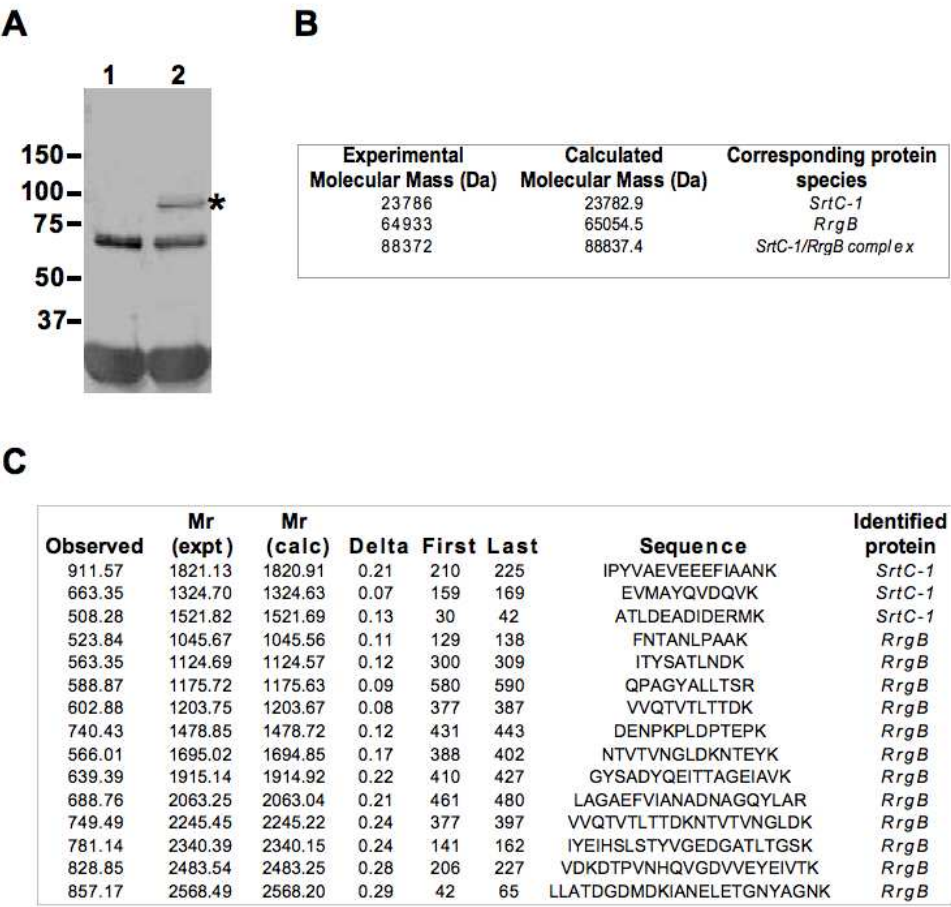


C



198x273mm (72 x 72 DPI)

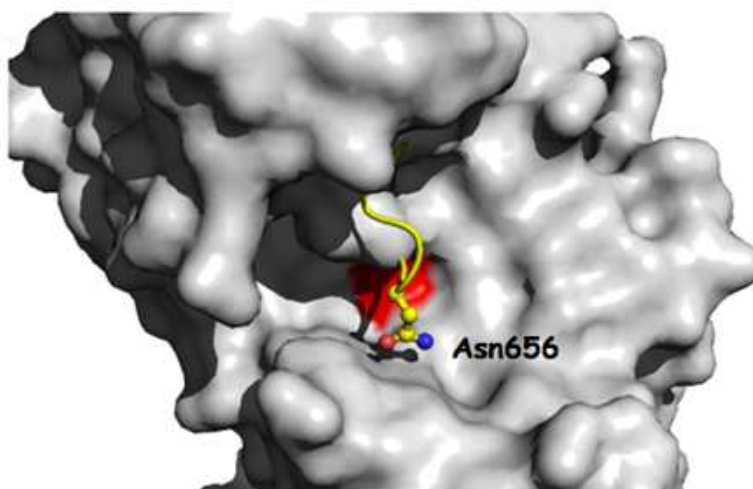
Figure 4



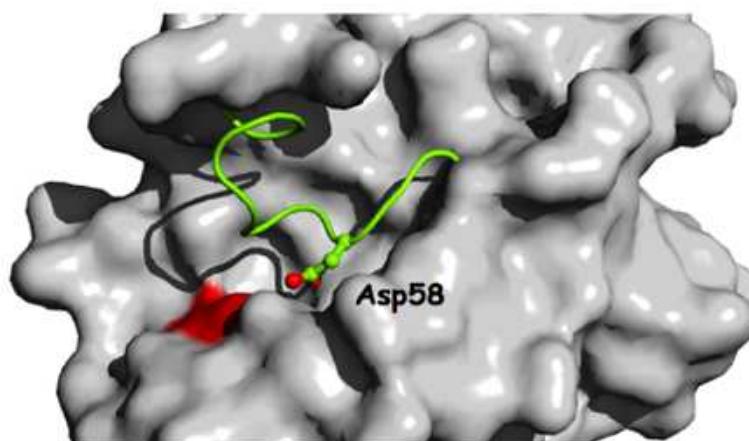
270x270mm (72 x 72 DPI)

Figure 5

A



B



JMB

Available online at www.sciencedirect.com


ScienceDirect

ELSEVIER

Two Crystal Structures of Pneumococcal Pilus Sortase C Provide Novel Insights into Catalysis and Substrate Specificity

Fabrice Neiers^{1,2,3}, Chaithanya Madhurantakam^{2†}, Stefan Fälker^{1,3‡}, Clothilde Manzano⁴, Andrea Dessen⁴, Staffan Normark¹, Birgitta Henriques-Normark^{1,3*‡} and Adnane Achour^{2*‡}

¹Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, SE-171 77 Solna, Sweden

²Center for Infectious Medicine, F59, Department of Medicine Huddinge, Karolinska University Hospital Huddinge, Karolinska Institutet, SE-141 86 Stockholm, Sweden

³Department of Bacteriology, Swedish Institute for Infectious Disease Control, SE-171 82 Solna, Sweden

⁴Institut de Biologie Structurale Jean-Pierre Ebel, UMR 5075 (CEA, CNRS, UJF, PSB), 41 rue Jules Horowitz, F-38027 Grenoble, France

Received 25 June 2009;
received in revised form
24 August 2009;
accepted 25 August 2009

The respiratory tract pathogen *Streptococcus pneumoniae* is a primary cause of morbidity and mortality worldwide. Pili enhance initial adhesion as well as the capacity of pneumococci to cause pneumonia and bacteremia. Pilus-associated sortases (SrtB, SrtC, and SrtD) are involved in the biogenesis of pneumococcal pili, composed of repeating units of RrgB that create the stalk to which the RrgA adhesin and the preferential pilus tip subunit RrgC are covalently associated. Using single sortase-expressing strains, we demonstrate that both pilin-polymerizing sortases SrtB and SrtC can covalently link pili to the peptidoglycan cell wall, a property shared with the non-pilus-polymerizing enzyme SrtD and the housekeeping sortase SrtA. Comparative analysis of the crystal structures of *S. pneumoniae* SrtC and SrtB revealed structural differences explaining the incapacity of SrtC, but not of SrtB, to incorporate RrgC into the pilus. Accordingly, site-directed mutagenesis of Thr¹⁶⁰ in SrtB to an arginine as in SrtC (Arg¹⁶⁰) partially converted its substrate specificity into that of SrtC. Solving two crystal structures for SrtC suggests that an opening of a flexible lid and a concomitant cysteine rotation are important for catalysis and the activation of the catalytic cysteine of pilus-associated sortases.

© 2009 Elsevier Ltd. All rights reserved.

Edited by R. Huber

Keywords: sortase; crystal structure; substrate recognition; catalysis; pili

Introduction

The human-specific pathogen *Streptococcus pneumoniae* (pneumococcus) is a primary cause of morbidity and mortality worldwide and represents

one of the four major infectious disease killers, together with HIV, malaria, and tuberculosis.^{1,2} *S. pneumoniae* is the main cause of respiratory tract infections such as otitis media, sinusitis, and community-acquired pneumonia, and it is also an important pathogen in invasive diseases such as septicemia and meningitis. Invasive pneumococcal disease is preceded by initial colonization of the nasopharynx, but the detailed mechanisms of adhesion are still not well understood.

Most bacterial pathogens have long filamentous structures known as pili or fimbriae extending from their surface. These structures are often involved in the initial adhesion of the bacteria to host cells and

*Corresponding authors. Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, SE-171 77 Solna, Sweden.

E-mail addresses: birgitta.henriques@smi.se; adnane.achour@ki.se.

† C.M. and S.F. contributed equally to this work.

‡ B.H.-N. and A.A. are shared last authors.

tissues, as well as to other bacteria, during colonization. Pilus expression mediates host cell responses that depend on the host receptor being recognized and may also increase the host inflammatory responses.^{3,4} Until recently, pili had only been recognized as an attribute of Gram-negative bacteria in which they have been well characterized from the genetic, biochemical, and structural perspectives. During the last few years, some major human Gram-positive bacterial pathogens have also been shown to encode pili, and their biogenesis, structure, and function are now being investigated.^{3,5-9}

Multiple-drug resistant *S. pneumoniae* strains have emerged over the last 30 years, requiring rapid development of novel preventive and therapeutic approaches including the development of protein-based vaccines.¹⁰ We have previously demonstrated that globally spreading penicillin nonsusceptible pneumococci belong to clonal lineages that frequently carry pili, encoded by the *rlrA* pathogenicity islet, suggesting that pili contribute to the successful spread of these clones.¹¹ Pili enhance initial adhesion as well as the subsequent capacity of the bacteria to cause pneumonia and bacteremia.³ The *rlrA* pilus islet in *S. pneumoniae* encodes three surface proteins, RrgA, RrgB, and RrgC, which form the pneumococcal pilus as well as three pilus-associated sortases, SrtB, SrtC, and SrtD.^{3,12,13} The stalk of the pneumococcal pilus is composed of repeating units of RrgB to which RrgA and RrgC are covalently associated. The RrgA protein also occurs associated to the cell wall.³ RrgC is frequently present at the

tip of pili but may also appear along the pilus stalk and in proximity of RrgA.¹² RrgA has been shown to be the major pilus-associated adhesin of pneumococcal pili, but RrgC may also have some binding functions.¹³

In general, sortases can be described as membrane-associated transpeptidases that transfer and covalently link surface proteins to the peptidoglycan cell wall.^{14,15} The linkage mechanism is initiated by the cleavage between the threonine and glycine residues of the conserved LPXTG motif (where X represents any amino acid) in substrate proteins.^{15,16} The C-terminal threonine of the surface protein is concomitantly linked to the sortase catalytic cysteine, after which it is incorporated to the cell wall through the formation of an amide bond between the C-terminal threonine and the stem peptide of the peptidoglycan.¹⁴ In contrast to these so-called housekeeping sortases, pilus-associated sortases covalently link the different subunits of the pilus to each other, resulting in the assembly of mature pili.^{5,8,17}

The relative contributions of pneumococcal-associated sortases, SrtB, SrtC, and SrtD, have recently been established.¹² While SrtD did not polymerize any of the Rrg subunits, SrtB and SrtC were redundant in terms of their activity, as both were able to polymerize RrgB in order to form the structural backbone of the pilus as well as link RrgA to this backbone. However, in contrast to SrtB, SrtC cannot incorporate RrgC to the pilus.¹² SrtC also differs from SrtB and SrtD in that the absence of

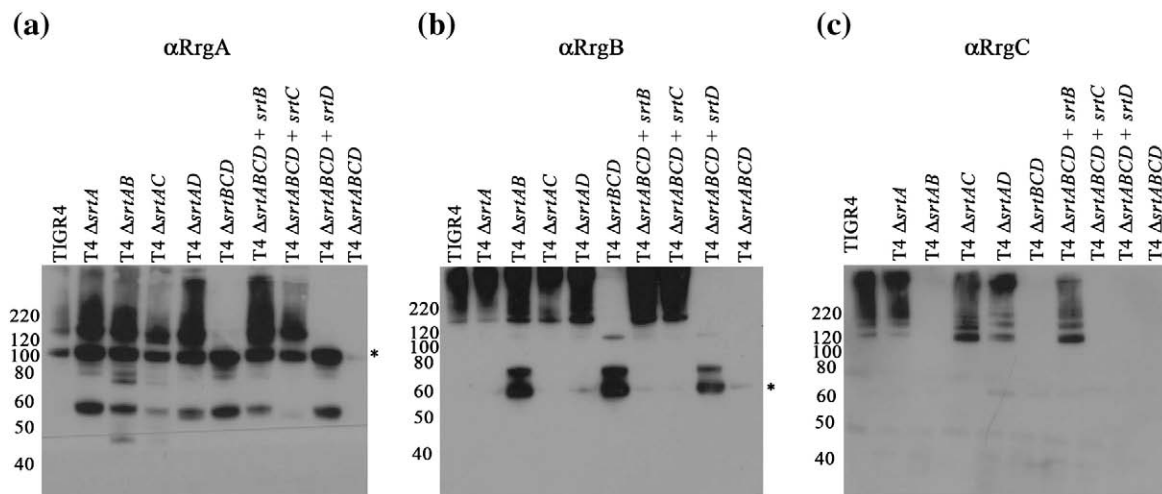


Fig. 1. Role of individual sortases in pilus biogenesis and cell wall anchoring. The production of polymeric high-molecular-mass cell wall-associated pili in isogenic T4 mutants was evaluated by immunoblotting for RrgA (a), RrgB (b), and RrgC (c). In all cases, cell wall proteins were separated by gradient SDS-PAGE, transferred to polyvinylidene fluoride, and probed. Approximate molecular masses are indicated in kilodaltons to the left. The three images show immunoblotting for RrgA (a), RrgB (b), and RrgC (c) in the wild-type TIGR4 ('T4'); the single sortase mutant strain T4Δ*srtA* ('Δ*srtA*'); the double-mutant strains T4Δ*srtAB* ('Δ*srtAB*'), T4Δ*srtAC* ('Δ*srtAC*'), and T4Δ*srtAD* ('Δ*srtAD*'); the triple-mutant strain T4Δ*srtBCD* ('Δ*srtBCD*') that expresses only SrtA; the *trans*-complemented *srtABCD* quadruple-mutant strains T4Δ*srtABCD*+*lacE::srtB* ('Δ*srtABCD*+*srtB*'), T4Δ*srtABCD*+*lacE::srtC* ('Δ*srtABCD*+*srtC*'), and T4Δ*srtABCD*+*lacE::srtD* ('Δ*srtABCD*+*srtD*'); and the quadruple-mutant strain T4Δ*srtABCD*. The predicted molecular masses for the RrgA (90 kDa) and RrgB (67 kDa) monomers are indicated by an asterisk. All four sortases are redundant in their capacity to transfer pili and/or pilin monomers to the cell wall. SrtD shows an *in vivo* activity. SrtB and SrtC differ only in their capacity to incorporate RrgC into the pilus.

either of the two latter enzymes, but not of SrtC, abrogates the focal presentation of pilus antigens at the bacterial cell surface.¹²

In the present study, we made use of single sortase-expressing strains to assess the capacity of each of the *S. pneumoniae*-associated sortases (SrtB, SrtC, and SrtD as well as the housekeeping sortase SrtA) to incorporate and link (or not) pilin monomers as well as the polymerized pilus to the peptidoglycan cell wall. We demonstrate that the pilus-polymerizing enzymes SrtB and SrtC can individually covalently anchor pilus polymers to the cell wall, reducing the substrate differences between these two sortases to their capacity to incorporate (SrtB) or not (SrtC) the RrgC pilin subunit into the pilus. Comparing the two crystal structures of the pilus-polymerizing sortase SrtC, solved in this study, with the previously described three-dimensional structure of SrtB,¹⁸ which exhibits different substrate specificities, as well as the nonpolymerizing enzyme (but with pilin cell wall anchoring activity) SrtD provides novel insights into the mechanisms underlying the activation of the catalytic cysteine Cys²²¹.

Results

All three pilus-associated sortases of *S. pneumoniae* can recognize the peptidoglycan cell wall as substrate for pili and pilin monomers

The first two steps within the proposed model of pilus assembly in Gram-positive organisms consist of the covalent linking of individual pilin subunits followed by the linkage of the polymerized pilus to the peptidoglycan cell wall.¹⁹ It has been demonstrated in *C. diphtheriae* that housekeeping and pilus-associated sortases are able to transfer pili to the cell wall.²⁰ Furthermore, recent studies also indicated that the capacity of a *S. pneumoniae* strain lacking the housekeeping sortase SrtA to link pili to peptidoglycan is not affected.²¹ We have demonstrated that a *S. pneumoniae* strain expressing only SrtA is able to transfer RrgA monomers to the cell wall.¹² Therefore, we hypothesized that at least one of the pilus-associated sortases and the housekeeping SrtA are redundant in their capacity to covalently link pilin proteins to the cell wall peptidoglycan.

In order to address this question and to further determine the specificities of SrtC as compared to SrtA, SrtB, and SrtD, pneumococcal strains lacking diverse combinations of the four sortases were constructed and analyzed for the presence of pili and/or pilus subunits in the cell wall fractions. Our results demonstrate that strains lacking SrtA or lacking a combination of SrtA and any of the pilus-associated sortases still contain pili in the cell wall fractions (Fig. 1). These data indicate that more than one pilus-associated sortase is able to catalyze the transfer of pilin proteins to the peptidoglycan. Surprisingly, even strains that express only one

pneumococcal sortase in the genome still display cell wall anchoring activity specific for pilin proteins. Indeed, both SrtB and SrtC transfer pili to the peptidoglycan, while SrtA and SrtD lacking pilus polymerization activity still link RrgA monomers and RrgB monomers (approximately 90 and 67 kDa, respectively, after processing) to the cell wall (Fig. 1). However, linkage of RrgB monomers to the cell wall seems to occur only if no or insufficient RrgB polymerization activity is present. As previously described,¹² an additional band corresponding to a processed RrgA protein with a smaller molecular mass (around 60 kDa) is observed in all strains except for the wild type and the strain without any sortase. This processing event remains to be explained. As also previously described,¹² only SrtB is able to incorporate RrgC in pili (Fig. 1). Furthermore, RrgC monomers are not transferred to the cell wall by any sortase, suggesting that cell surface exposure of this protein is strictly dependent on SrtB-mediated linkage to either RrgA or RrgB. A

Table 1. Data collection and refinement statistics

	SrtC (PDB code 3G66)	SrtC (PDB code 3G69)
<i>Data collection</i>		
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	<i>a</i> = 48.9, <i>b</i> = 96.9, <i>c</i> = 98.9	<i>a</i> = 48.6, <i>b</i> = 96.5, <i>c</i> = 98.8
α , β , γ (°)	α = β = γ = 90	α = β = γ = 90
Observed reflections	350,859 (46,799)	228,540 (28,510)
Unique reflections	52,404 (7580)	31,946 (4490)
Redundancy	6.7 (6.3)	7.2 (6.3)
Completeness (%)	99.8 (99.2)	99.2 (97.2)
R_{merge}^a (%)	9.3 (51.7)	10.9 (30.5)
$I/\sigma(I)$	14.6 (3.4)	16.1 (6.3)
<i>Refinement</i>		
Resolution (Å)	32.3–1.7 (1.74–1.70)	43.6–2.0 (2.05–2.00)
No. reflections	49,719	30,362
$R_{\text{work}}/R_{\text{free}}^b$ (%)	18.7/23.5	19.2/25.8
No. atoms		
Protein	3145	3126
Mes/ion	24	34
Water	557	467
<i>B</i> -factors (Å ²)		
Protein	20.3	23.7
Mes	25.6	24.2
Sulfate ion	—	37.8
Water	40.5	36.2
R.m.s. deviations		
Bond lengths (Å)	0.015	0.023
Bond angles (°)	1.74	2.13
Ramachandran plot (%)	86.8, 11.1, (favored, allowed, generous, disallowed)	87.2, 10.4, 2.4, 0
PDB code	3G66	3G69

Values in parentheses are for highest-resolution shell.

^a $R_{\text{merge}} = \sum_i \sum_{hkl} |I_i(hkl) - \langle I(hkl) \rangle| / \sum_i \sum_{hkl} I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl) \rangle$ is the weighted average intensity for all observations *i* of reflection *hkl*.

^b R_{cryst} and $R_{\text{free}} = (\sum_i |F_o| - |F_c|) / (\sum_i |F_o|)$, where $|F_o|$ is the observed structure-factor amplitude and $|F_c|$ is the calculated structure-factor amplitude. R_{free} is based on 5% of the total reflections excluded from refinement.

quadruple-mutant strain, which does not express any of the pneumococcal sortases, does not contain any pilin protein in the cell wall, demonstrating that covalent linking of pili and/or pilin monomers to the peptidoglycan is strictly sortase dependent and that no other factor/molecule can accomplish this process (Fig. 1). Our mutation study provides the first evidence for an *in vivo* enzymatic role of SrtD in pilus biogenesis in that it is able to covalently link RrgA and RrgB proteins to the peptidoglycan. This enzymatic activity of SrtD probably explains its requirement for regular focal presentation of pilus antigens on the bacterial surface.¹² On the basis of these results, the comparison of the crystal structures

of SrtC, SrtB, and SrtD should highlight conserved features characteristic of pilus-specific sortases as well as allow the identification of key residues that dictate differences in substrate specificity.

Overall structure of sortase C

Two crystallization conditions were determined for SrtC.²² The two crystal structures were solved to 1.7- and 2.0-Å resolutions, respectively (Table 1). The two SrtC structures are identical except for sections of a loop encompassing residues 79–100 that form a lid, as well as for the conformation of the key catalytic residue Cys²²¹. The fold of SrtC corresponds

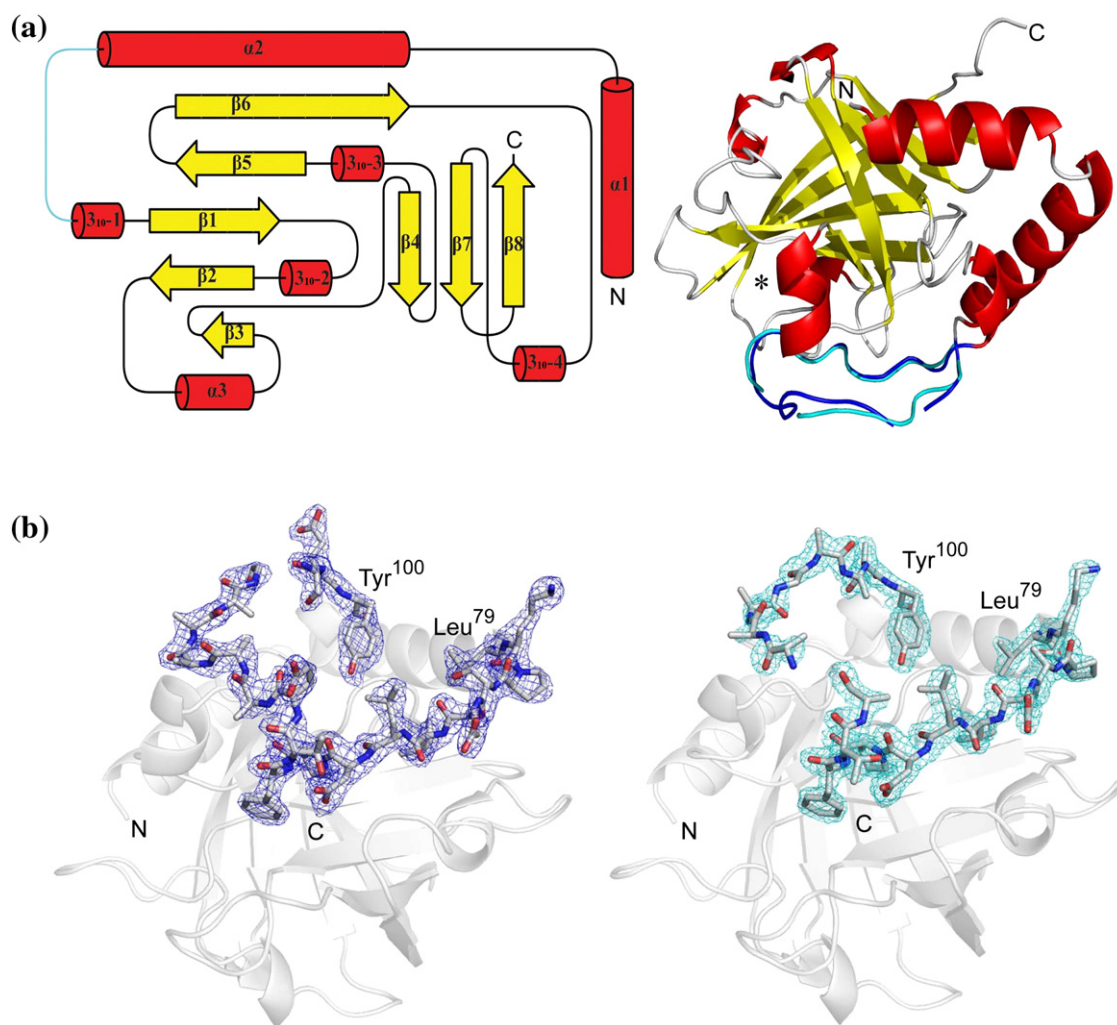


Fig. 2. Topology diagram and overall view of the pilus-associated SrtC from *S. pneumoniae*. (a) The secondary structure diagram of SrtC is displayed to the left with the β -strands in yellow and the 3_{10} - and α -helices in red. The N- and C-termini of the enzyme are indicated by N and C, respectively. The lid specific to pilus-associated sortases is in cyan. The figure was generated using the program TopDraw.²³ An overall view of the crystal structures of the *S. pneumoniae*-derived SrtC is presented to the right. The yellow-colored β -strands form the β -barrel core of the enzyme. Helices are colored in red. The loop consisting of residues 79–100 and that links $\alpha 2$ -helix and $\beta 1$ -strand forms a flexible lid that closes the active site, and has different conformations in the two crystal structures of SrtC presented in this study. The loops found in the crystal structures of SrtC (PDB code 3G66) and SrtC (PDB code 3G69) are colored cyan and blue, respectively. The N- and C-termini of the enzyme are indicated by N and C, respectively. (b) The $2F_o - F_c$ electron densities (contoured at 1.0σ) of the lid regions are displayed in blue for SrtC to the left (PDB code 3G69) and in cyan for SrtC (PDB code 3G66) to the right. Residues 79 and 100 are indicated in the figure. Residues 79–100 that correspond to the lid regions of SrtC are shown in stick mode, while the rest of the two structures are colored in transparent white.

to the previously described β -barrel structure, specific for sortases, composed of eight anti-parallel β -strands linked by multiple α - and 3_{10} -helices (Figs. 2a and 3).^{24–28} However, in contrast to house-keeping sortases such as SrtA,²⁴ the structure of SrtC contains a lid that is common to pilus-associated sortases such as SrtB and SrtD in *S. pneumoniae*.¹⁸

As expected from the sequence identity (58%), the overall fold of SrtC is similar to the recently described crystal structure of SrtB,¹⁸ with an average root mean square deviation (rmsd) upon superposition of all C $^{\alpha}$ atoms for the two structures of 1.2 Å (Fig. 4). Despite lower sequence identity (27%), SrtC also resembles SrtD¹⁸ (with an rmsd of 1.9 Å), although major localized differences can be noted such as the presence of an additional C-terminal helix in SrtD (Figs. 3 and 4). The active site is localized on the external face of the β -barrel and is closed by the lid (Fig. 2a). Although the lid feature is shared by SrtB and SrtD, its conformation is

different. The relative mobility is evidenced by the poorer electron density for most of the side chains and some of the main chain of the residues forming the lid (residues 79–85 and 89–100) (Fig. 2b). In contrast, the electron density of the central section of the lid (residues 86–88) that is anchored within the active site through several interactions with key catalytic residues is well defined (Fig. 2a and b). The relatively higher *B*-factor values observed for the lid regions in both crystal structures (41.2 and 45.7 Å² on average for residues 79–100 in 3G66 and 3G69, respectively) when compared to the overall values (20.3 and 23.7 Å² on average for 3G66 and 3G69, respectively) argue for a higher mobility of the lid. The quality of the electron density allows a well-defined positioning of most of the main-chain residues of the lid. Thus, the main chain of the lid can adopt at least two different conformations, reflecting further the mobility of the lid (Fig. 2a and b).

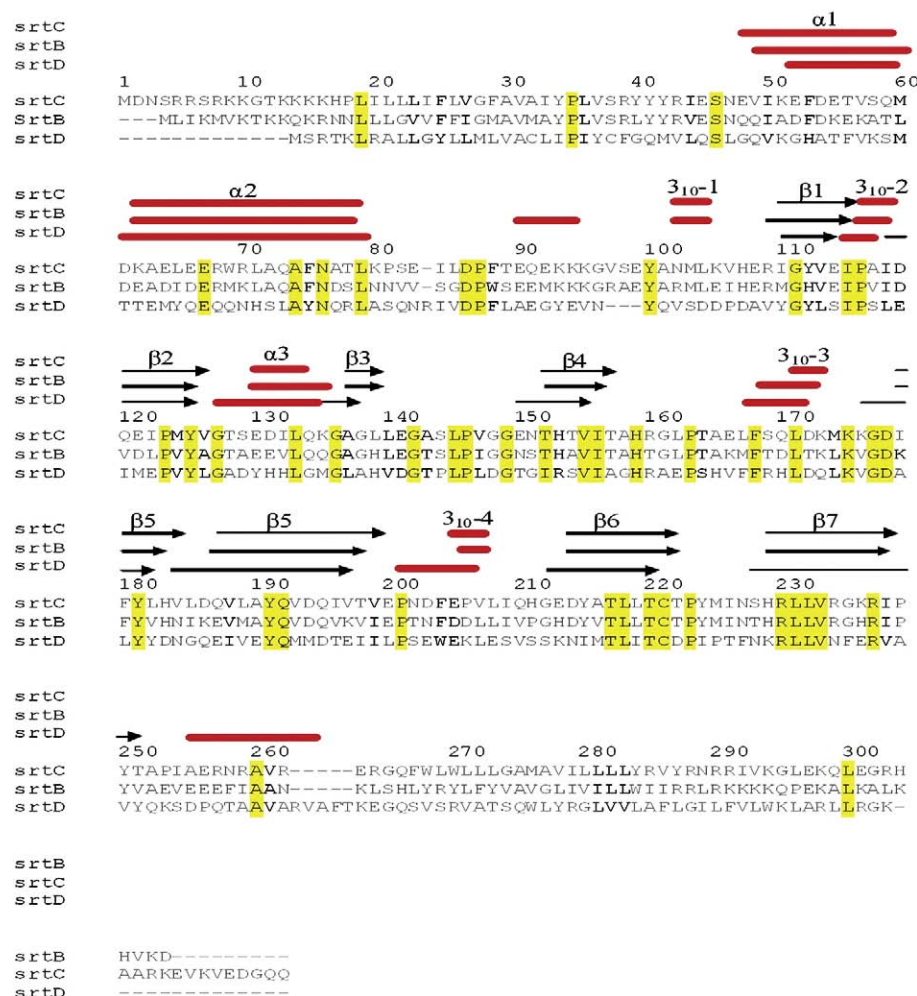


Fig. 3. Sequence alignment of the pilus-associated sortases SrtC, SrtB, and SrtD from *S. pneumoniae*. The numbering used for SrtB and SrtD is aligned on the numbering of the sequence of the SrtC. The secondary structures are indicated for each sortase with black arrows (β -strands) and red lines (3_{10} - and α -helices). Identical and similar residues are indicated with black letters on yellow background and black letters only, respectively. The sequences were aligned using the program Bioedit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>).

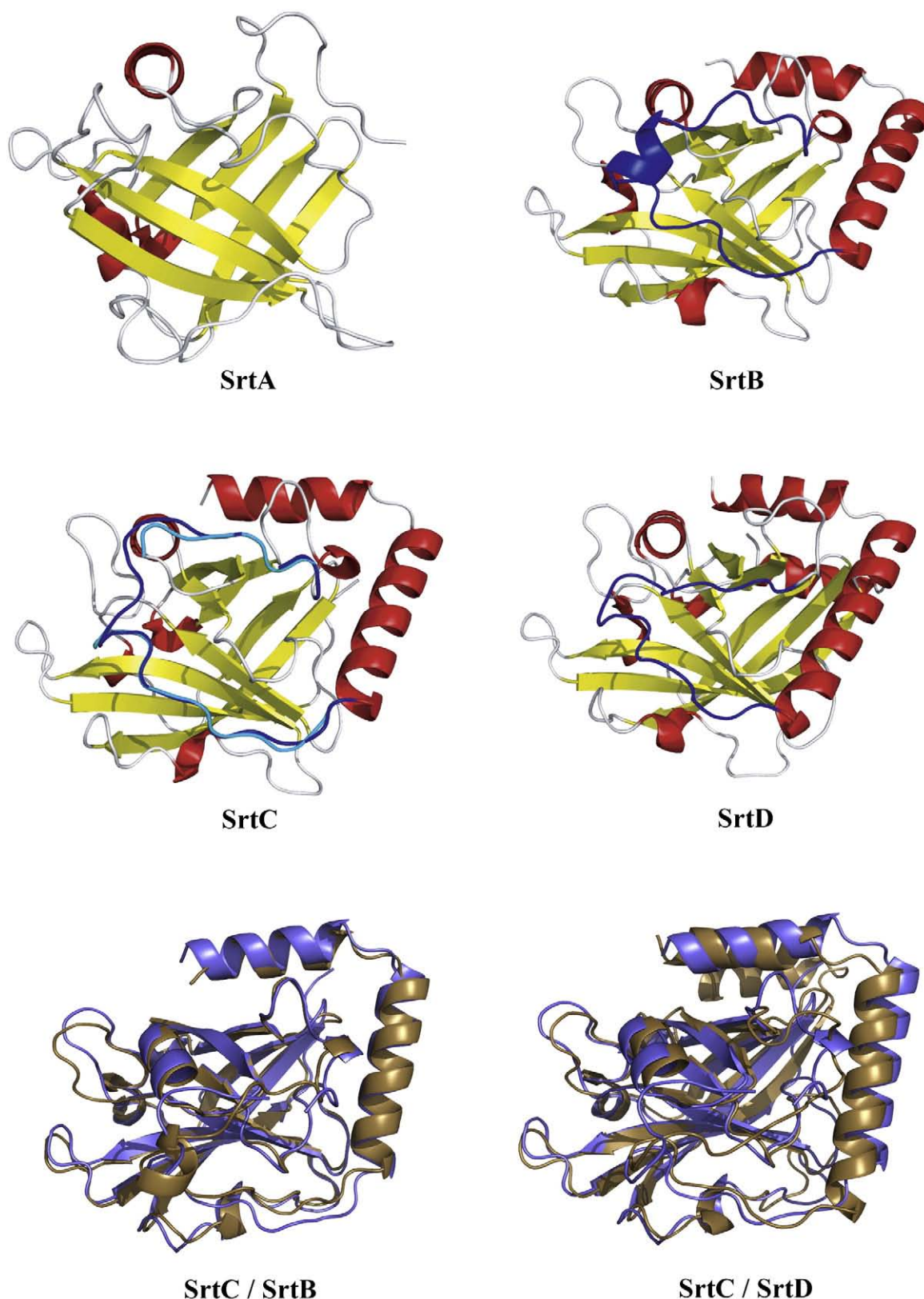


Fig. 4. All pilus-associated sortases from *S. pneumoniae* contain a lid that covers the active site. Structural comparison of the three pilus-associated *S. pneumoniae*-derived sortases, SrtB (PDB code 2W1J), SrtC, and SrtD (PDB code 2W1K), and the housekeeping sortase SrtA derived from *Staphylococcus aureus* (PDB code 1T2P). All three pilus-associated sortases SrtB, SrtC, and SrtD from *S. pneumoniae* display an additional loop (blue) that closes the active site that is absent in the housekeeping sortase SrtA from *S. aureus*. The superpositions of SrtC (purple) with SrtB (gold) and SrtD (gold) are displayed at the bottom of the figure. The structures differ by the lid positioning and an additional helix in the C-terminal region that is only present in SrtD.

The tip of the lid is anchored within the active site through multiple interactions with key catalytic residues

The carboxylate group of the side chain of Asp⁸⁶ on the tip of the lid forms a salt bridge with a 'fork-stick' geometry with the side chain of Arg²³⁰ within the active site (Fig. 5).²⁹ The sulfur atom of the catalytic residue Cys²²¹ forms a sulfur-aromatic interaction with the aromatic ring of Phe⁸⁸, strengthening the anchoring of the lid within the active site.³⁰ Sulfur-aromatic interactions between cysteine and phenylalanine residues have been described previously as a combination of hydrophobic and electrostatic interactions.³¹ Interestingly, sulfur-aromatic interactions are also formed between the catalytic cysteine and a tryptophan or a phenylalanine residue on the lids of SrtB and SrtD, respectively.¹⁸ This indicates a general mechanism for SrtC, SrtB, and SrtD in which

a sulfur-aromatic interaction between Cys²²¹ and an aromatic residue on the lid results in its anchoring within the active site.

Comparative structural analysis suggests a molecular mechanism for catalysis within the active site

It has previously been demonstrated that Cys²²¹ is essential for the catalytic activity of housekeeping and pilus-associated sortases.^{12,32} However, the molecular mechanisms leading to the deprotonation of Cys²²¹ remain to be defined. The crystal structures of SrtC indicate that the distance between residues His¹⁵⁹ and Cys²²¹ (>5.1 Å) is not compatible with a direct deprotonation *via* a thiolate-imidazolium pair. Omit map analysis of SrtC (PDB code 3G66) demonstrates that Cys²²¹ can adopt two different conformations within the active site of

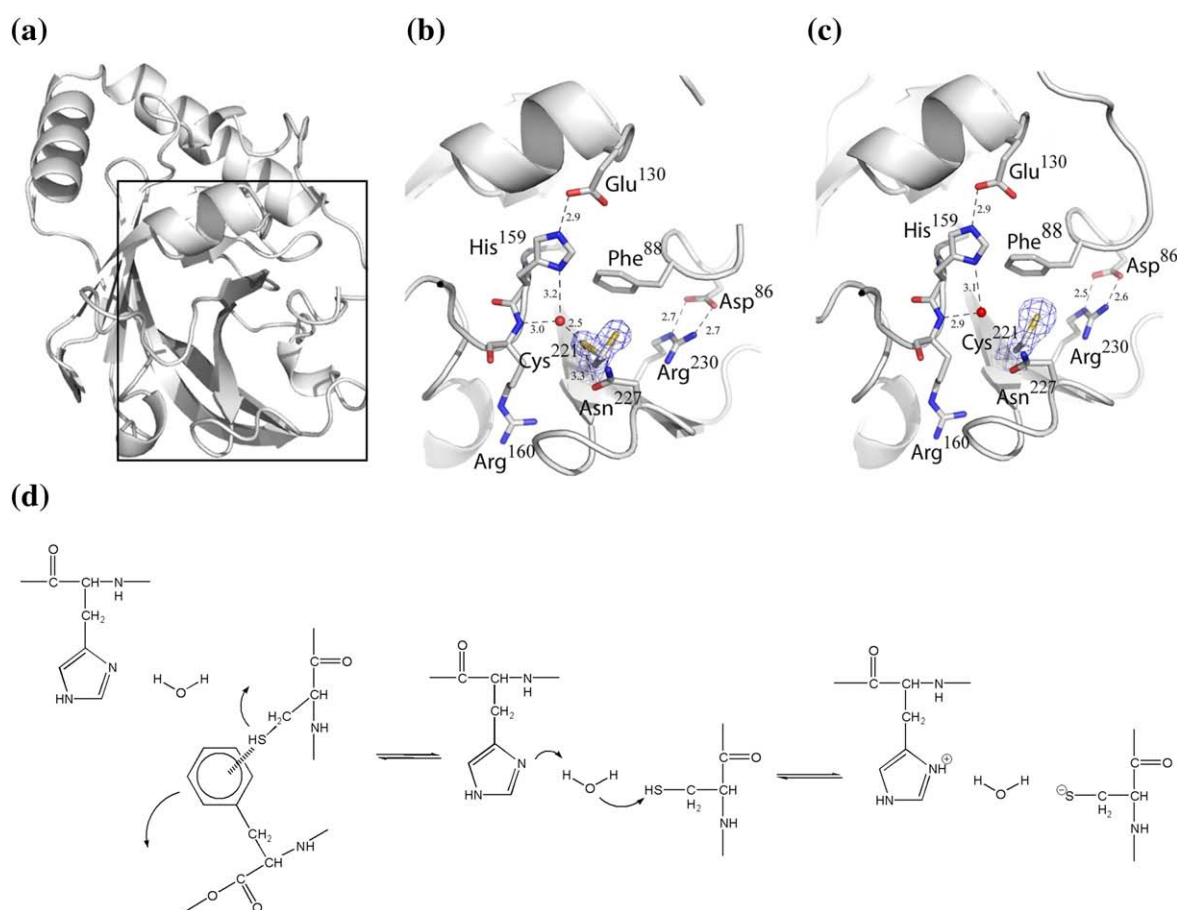


Fig. 5. The catalytic cysteine residue Cys²²¹ can take two different conformations within the active site of SrtC. (a) An overview of the crystal structure of SrtC; the square indicates the region that is enlarged in (b) and (c). The active sites of SrtC are presented in (b) (PDB code 3G66) and (c) (PDB code 3G69), respectively. While the catalytic cysteine Cys²²¹ adopts two different conformations (modeled with 0.5 occupancy) in (b), the same residue adopts only one conformation in (c). Electron density (contoured at 1.2 Å) is displayed in blue. Water molecules are indicated by red spheres and hydrogen bonds are shown as broken lines with distances indicated in angstrom. The figure was generated using the program PyMol (Delano Scientific LLC). In (d), a mechanism for the activation of the catalytic cysteine in the pilus-associated sortases is proposed. The substrate binding can be the driving force that disrupts the stabilizing interaction between the catalytic cysteine and the aromatic residue of the loop (Phe for SrtD and Trp for SrtB) as observed in the structure display in (c). The cysteine residue becomes available for proton transfer *via* a conserved water molecule (similar to the structure shown in Fig. 4b), resulting in the formation of the active thiolate form.

SrtC (Fig. 5b and c). One conformation of Cys²²¹ forms sulfur-aromatic interactions with Phe⁸⁸, while the second conformation forms a hydrogen bond interaction with a water molecule (WAT⁷⁴). It should be noted that a water molecule is localized at exactly the same position in the crystal structures of all three described pilus-associated sortases. A network of hydrogen bond interactions is formed within this section of the active site, reaching from Cys²²¹ to Glu¹³⁰. In addition to the hydrogen bond interaction formed with Cys²²¹, WAT⁷⁴ also forms two hydrogen bond interactions with His¹⁵⁹ and Arg¹⁶⁰. In turn, the nitrogen atom N^ε of His¹⁵⁹ forms a hydrogen bond interaction with Glu¹³⁰ (Fig. 5b).

Omit map analysis clearly shows that only one conformation is present for Cys²²¹ in the second crystal structure of SrtC (PDB code 3G69) (Fig. 5c). It should be noted that the position of the water molecule WAT⁷⁴ is conserved in both crystal structures of SrtC (Figs. 4c and 5b). Thus, other factors may result in the dual or single conformations of Cys²²¹ observed in the two crystal structures of SrtC. The interaction with a substrate could open the lid of SrtC, resulting in the disruption of sulfur-aromatic interactions between Phe⁸⁸ and Cys²²¹ and an increased propensity for the second conformation of Cys²²¹ to occur. The water molecule WAT⁷⁴ would then be able to transfer the proton from Cys²²¹ to His¹⁵⁹, which is stabilized and properly oriented through hydrogen bond interactions with Glu¹³⁰. The interactions with Glu¹³⁰ could thereafter result in an increase of the pK_a of His¹⁵⁹, allowing for its protonation (Fig. 5d). Cysteine proteases have been previously described to perform a nucleophilic attack on the carbonyl carbon at the scissile peptide bond of the substrate by the thiolate form of Cys²²¹,

resulting in the formation of a thioacyl intermediate. A similar relay between the nucleophilic group and the general base by a water molecule has been proposed before for β -lactamases.³³ As also suggested by studies in housekeeping sortases,²⁵ the oxyanion formed on the thioacyl intermediate would probably be stabilized by interactions with Arg²³⁰, which is conserved in all sortases. The second step of the catalytic mechanism would lead to the release of the cleaved peptide before the attack of the thioacyl intermediate by an amine group (from the side chain of the lysine residue in the case of the pilin motif), resulting in the production of a cross-linked peptide and the regeneration of the enzyme. Future determination of the thioacyl intermediate of pilus-associated sortases should provide a better understanding of this mechanism.

Structural differences between SrtC and SrtB explain the incapacity of SrtC to incorporate RrgC in the pilus

In order to build pili, pilus-associated sortases cut LPXTG-like motifs in Rrg molecules (YPRTG for RrgA, IPQTG for RrgB, and VPDTG for RrgC). The cleaved Rrg molecule is thereafter linked to a pili motif such as YPKN, particularly in the case of RrgB–RrgB linking.⁸ Subtle modifications within the active sites of SrtC and SrtB may explain the observed differences in the fine specificity and the recognition of substrates by these two sortases.

While the active site of SrtB is mainly occupied by the lid, surface analysis indicates the presence of an empty pocket that could be used for substrate recognition (Fig. 6a). Importantly, the corresponding pocket is occupied by the side chain of residue

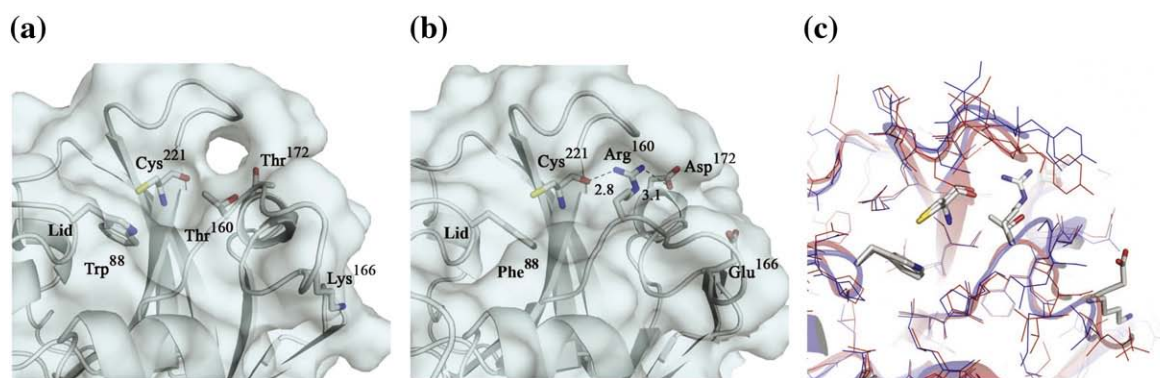


Fig. 6. Comparative structural analysis suggests a molecular explanation for the substrate specificity. Comparison of the active sites of SrtB (PDB code 2W1J) (a) and SrtC (b) provides a structural explanation for the functional differences observed between these two sortases. The side chains of the catalytic cysteines Cys²²¹, as well as the side chains of residues within and near the vicinity of the active site, are displayed. The side chain of the residue Phe⁸⁸ localized on the lid of SrtC, as well as the side chain of the corresponding residue Trp⁶⁰ in SrtB, is displayed. The side chain of residue Arg¹⁶⁰ in SrtC forms hydrogen bond interactions with the side chain of the Asp¹⁷² and the oxygen atom of the main chain of Cys²²¹. Hydrogen bonds are shown as black broken lines with distances indicated in angstrom. The side chains of residues Thr¹⁶⁰ and Thr¹⁷² in SrtB that correspond to Arg¹⁶⁰ and Asp¹⁷² in SrtC, respectively, are smaller, creating a substrate pocket in the active site of SrtB that is not present in SrtC. It should also be noted that the negatively charged residue Glu¹⁶⁶ present at the surface of SrtC within the vicinity of the active site is substituted to a positively charged lysine residue Lys¹⁶⁶ in SrtB. A superposition of SrtC (in red) and SrtB (in blue) is provided in (c). As indicated in (a) and (b), the majority of the residues within the active site are conserved (represented in line mode) except for a few residues (represented in stick mode).

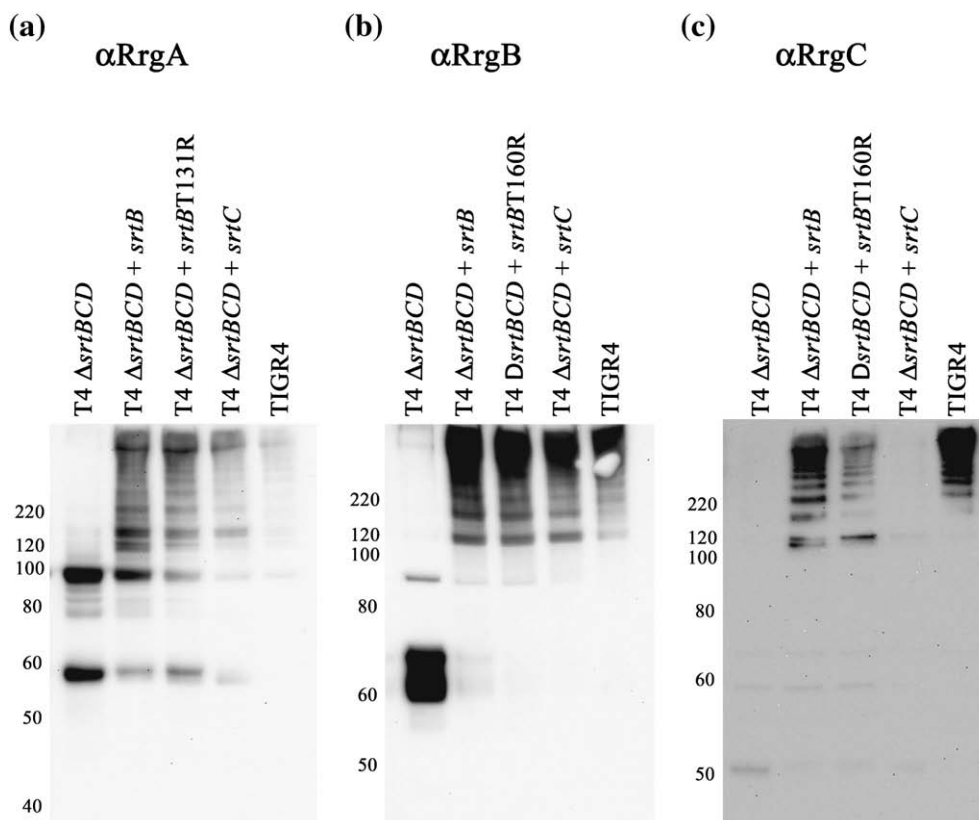


Fig. 7. Substitution of Thr¹⁶⁰ in SrtB to an arginine (as in SrtC) impairs its capacity to incorporate RrgC. Production of polymeric high-molecular-mass cell wall-associated pili in isogenic T4 mutants was evaluated by immunoblotting for RrgA (a), RrgB (b), and RrgC (c). Approximate molecular masses in kilodaltons are indicated to the left. The three images show immunoblotting for RrgA, RrgB, and RrgC in the triple-mutant strain T4Δ*srtBCD* ('Δ*srtBCD*'), the *trans*-complemented *srtBCD* triple-mutant strain T4Δ*srtBCD*+*lacE::srtB* ('Δ*srtBCD*+*srtB*'), the *srtBCD* triple-mutant strain *trans*-complemented with SrtB with the point mutation Thr¹⁶⁰Arg T4Δ*srtBCD*+*lacE::srtBT160R* ('Δ*srtBCD*+*srtB*'T160R), and the *trans*-complemented *srtBCD* triple-mutant strain T4Δ*srtBCD*+*lacE::srtC* ('Δ*srtBCD*+*srtC*'). Substitution of Thr¹⁶⁰, one of the few positions in SrtB differing with SrtC, to an arginine as in SrtC affects the capacity of the mutated SrtB to incorporate RrgC.

Arg¹⁶⁰ in the crystal structures of SrtC (Fig. 6b). The side chain of Arg¹⁶⁰ in SrtC is stabilized through double hydrogen bond interactions with the main chain of Cys²²¹ and the side chain of Asp¹⁷². Conversely, the positions corresponding to Arg¹⁶⁰ and Asp¹⁷² in SrtC are occupied by two threonines (Thr¹⁶⁰ and Thr¹⁷²) in SrtB (Fig. 6a and b). Indeed, repeated *in vivo* analysis of the effects of the substitution of Thr¹⁶⁰ to an arginine in SrtB clearly reduces its capacity to incorporate RrgC in the pilus, while its capacity to incorporate both RrgA and RrgB remains unaffected (Fig. 7). While the observed effect reveals the importance of Thr¹⁶⁰ in SrtB, it also indicates that other residues such as Thr¹⁷² may play a role in the specificity of SrtB in incorporating RrgC (Fig. 6b).

Discussion

The three-dimensional structures of pneumococcal SrtC provide novel information regarding the catalysis and the specificity of this enzyme. The catalytic cysteine of housekeeping sortases such as

the *Staphylococcus aureus*-derived SrtA is localized within a hydrophobic pocket,²⁴ which can probably protect against erroneous oxidation. In contrast, the corresponding catalytic cysteine is hidden by a lid in the pilus-associated SrtB, SrtC, and SrtD, indicating that unregulated oxidation is prevented through an alternative pathway. The two crystal structures of SrtC presented within this study also indicate that the lid is mobile and can take at least two different conformations, suggesting a second important role for the lid in the regulation of substrate accessibility to the active site.¹⁸ The presence of a lid, however, does not strictly correlate to pilin-pilin transpeptidase activity, since SrtD, having a lid, lacks this activity. Furthermore, it is not required for the enzyme to discriminate between peptidoglycan and pilin substrates, since SrtB and SrtC, both having a lid, can use both types of substrates. Also, SrtD, which contains a lid, and SrtA, which most likely lacks a lid in analogy to staphylococcal SrtA, can both use peptidoglycan as substrate for anchoring of pilin. Our structural analysis indicates the dual conformation of the catalytic cysteine in SrtC as an intrinsic general property of pilus-associated

sortases and suggests a general mechanism of catalysis for the pilus-associated sortases described here. Future pK_a determination and *in vitro* kinetic studies should provide a better understanding of this mechanism that is probably similar for all three pilus-associated sortases in pneumococci.

The *in vivo* experiments provided in this study demonstrate the dual functions of SrtC and SrtB comprising pilus building and LPXTG-cell wall anchoring activity. From an evolutionary point of view, our results suggest the existence of a common ancestor for pilus-associated and housekeeping sortases. Since the housekeeping function is conserved among most known Gram-positive bacterial organisms, we hypothesize that the common ancestor was a housekeeping sortase with a β -barrel structure fold. Pilus-associated sortases have most probably evolved from this common ancestor through the addition of structural features such as the lid and an additional α -helix in the case of SrtD.

In this context, it may not be surprising to note that the charge of the solvent-exposed lid is similar when comparing SrtB and SrtC. Both lids possess a positively charged portion composed of three lysine residues (positions 93–95), in contrast to the entire lid in SrtD that is exclusively composed of hydrophobic or negatively charged residues (Fig. 3). The lid could play an important role in the interaction of pilus-associated sortases with substrates, since it can open or close access to the active site. The observed differences in charge on the lid can represent another important aspect that explains why SrtB and SrtC act as pilus-polymerizing enzymes, in contrast to SrtD.

Comparative analysis indicates that very few residues differ between the active sites of SrtB and SrtC. The crystal structures of SrtC reveal that the side chain of residue Arg¹⁶⁰ in SrtC fills a substrate recognition pocket in the active site. Substitution of the corresponding residue Thr¹⁶⁰ in SrtB to an arginine demonstrates the essential role of this residue for the incorporation of RrgC. Interestingly, this pocket is also filled in SrtD by an arginine that forms hydrogen bond interactions with the catalytic cysteine and an aspartate residue. This may provide a structural explanation to the fact that SrtD is not able to incorporate RrgC into a pilus polymerized by SrtC. The observed reduction of the *in vivo* capacity of the SrtB mutated variant to incorporate RrgC also indicates that other residues can be involved in the nonrecognition of RrgC by SrtC. Asp¹⁷², which stabilizes the positioning of Arg¹⁶⁰ in SrtC *via* a hydrogen bond interaction, may play an important role for the closing of the recognition pocket in SrtB. Also, if we consider the size of the RrgC protein (30 kDa), the interaction might not be limited to the sortase active site, but other residues involved in surface charge could also determine RrgC recognition. The only difference in surface charge close to the active site is the positively charged Lys¹⁶⁶ in SrtB replaced by the negatively charged Glu¹⁶⁶ in SrtC (Fig. 6a and b).

The molecular mechanisms underlying the formation of the pilus still remain to be fully assessed. Our results demonstrate cell wall anchoring activity of SrtB, SrtC, as well as SrtD. Analysis of the two three-dimensional structures of SrtC suggests the mobility of the lid specific to pilus sortases as well as its importance for the activation of the catalytic cysteine residue. *In vivo* experiments demonstrate that the only functional difference between SrtB and SrtC is the incorporation of RrgC in the pilus. The structural comparison of SrtC and SrtB suggested the importance of specific amino acids in the vicinity of the active site for the observed functional differences between these two sortases. Finally, comparison of the two structures of SrtC suggests that rotation of the active-site cysteine close to a water molecule, in association with lid mobility, may allow for deprotonation of the cysteine, resulting in the thiolate form required for forming a thioacyl intermediate with the substrate.

Materials and Methods

Bacterial strains, media, and growth conditions

All the pneumococcal strains used within this study, including isogenic mutant derivatives, are described in Table 2. The insertion-deletion mutagenesis used for most strains as well as the strategy to achieve complementation *in trans* are described elsewhere.^{3,35} The primers used for each strain are indicated in Table 2, and all primers are described in Table 3. All resulting strains were checked by PCR, sequencing, and immunogenicity, thereby demonstrating a lack of polar effects of deletion of one subunit gene on the expression of other genes. Unless otherwise noted, bacteria were streaked from frozen stocks onto blood plates with appropriate selection for overnight growth and inoculated briefly into pre-warmed DS (dextrose-serum) medium (OXOID manual, 1990). DS was thereafter inoculated into pre-warmed C+Y medium in order to achieve an absorbance of 0.05 when measured at a wavelength of 620 nm. Cultures were permitted to grow to mid-log ($A_{620}=0.4$) at 37 °C without agitation before collection for experimentation.

Cell wall preparation and immunoblotting

Cell wall-associated proteins were isolated from genetically defined strains of *S. pneumoniae* and analyzed as previously described.³

Protein expression and purification

DNA encoding residues 46 to 267 of the SrtC protein from *S. pneumoniae* (TIGR4, SP_0467) was cloned into the NdeI and SacI restriction sites of the pET24c plasmid (Novagen). The resulting construct was transformed into *Escherichia coli* BL21 (DE3) pLysS (Novagen) for overexpression at 37 °C. The cultures were maintained at 37 °C for 12 h after induction with isopropyl-D-1-thiogalactopyranoside. SrtC was purified following previously described procedures.²²

Table 2. Characteristics of strains or mutants used in the study

Strain/mutant	Relevant characteristics	Primers employed in construction	Source/reference
T4	TIGR4 pilated invasive isolate	Not applicable	3
T4Δ <i>srtA</i>	Kan ^R , strain lacking <i>srtA</i>	Upstream: SrtA-1 and SrtA-2; downstream: SrtA-3 and SrtA-4; Kan-resistance cassette from Janus-fragment: Dam406 and Kan-Rev-Apa	34; this study
T4Δ <i>srtAB</i>	Erm ^R , Kan ^R , strain lacking <i>srtA</i> and <i>srtB</i>	SrtA-1/SrtA-4 fragment moved from T4Δ <i>srtA</i> to T4Δ <i>srtB</i>	12; this study
T4Δ <i>srtAC</i>	Erm ^R , Kan ^R , strain lacking <i>srtA</i> and <i>srtC</i>	SrtA-1/SrtA-4 fragment moved from T4Δ <i>srtA</i> to T4Δ <i>srtC</i>	12; this study
T4Δ <i>srtAD</i>	Erm ^R , Kan ^R , strain lacking <i>srtA</i> and <i>srtD</i>	SrtA-1/SrtA-4 fragment moved from T4Δ <i>srtA</i> to T4Δ <i>srtD</i>	12; this study
T4Δ <i>srtBCD</i>	Erm ^R , strain lacking <i>srtB</i> , <i>srtC</i> , and <i>srtD</i> , expressing <i>srtA</i>	Upstream: SrtB-1 and SrtB-2; downstream: SrtD-3 and SrtD-4	12
T4Δ <i>srtBCD</i> + <i>lacE::srtB</i>	Spc ^R , Erm ^R , T4Δ <i>srtBCD</i> containing <i>srtB</i> in trans	Upstream: NlacEF and SpcR1; coding sequence: SrtB-aad9 and Sp0479-SrtB; downstream: NlacER and Sp0479R	12
T4Δ <i>srtBCD</i> + <i>lacE::srtC</i>	Spc ^R , Erm ^R , T4Δ <i>srtBCD</i> containing <i>srtC</i> in trans	Upstream: NlacEF and SpcR1-EcoRI; coding sequence: SrtC5'-EcoRI and SrtC3'-HindIII; downstream: NlacER and Sp0479R-HindIII	12
T4Δ <i>srtBCD</i> + <i>lacE::srtBT131R</i>	Spc ^R , Erm ^R , T4Δ <i>srtBCD</i> containing <i>srtB</i> in trans, Thr ¹³¹ changed to Arg	Upstream: NlacEF and SrtBT131Rrev; downstream: NlacER and SrtBT131Rfor; fragments from T4Δ <i>srtBCD</i> + <i>lacE::srtB</i> fused by overhang extension PCR	This study
T4Δ <i>srtABCD</i>	Erm ^R , Kan ^R , strain lacking <i>srtA</i> , <i>srtB</i> , <i>srtC</i> and <i>srtD</i>	SrtA-1/SrtA-4 fragment moved from T4Δ <i>srtA</i> to T4Δ <i>srtBCD</i>	12; this study
T4Δ <i>srtABCD</i> + <i>lacE::srtB</i>	Spc ^R , Erm ^R , Kan ^R , quadruple-sortase mutant strain containing <i>srtB</i> in trans	SrtA-1/SrtA-4 fragment moved from T4Δ <i>srtA</i> to T4Δ <i>srtBCD</i> + <i>lacE::srtB</i>	12; this study
T4Δ <i>srtABCD</i> + <i>lacE::srtC</i>	Spc ^R , Erm ^R , Kan ^R , quadruple-sortase mutant strain containing <i>srtC</i> in trans	SrtA-1/SrtA-4 fragment moved from T4Δ <i>srtA</i> to T4Δ <i>srtBCD</i> + <i>lacE::srtC</i>	12; this study
T4Δ <i>srtABCD</i> + <i>lacE::srtD</i>	Spc ^R , Erm ^R , Kan ^R , quadruple-sortase mutant strain containing <i>srtD</i> in trans	SrtA-1/SrtA-4 fragment moved from T4Δ <i>srtA</i> to T4Δ <i>srtBCD</i> + <i>lacE::srtD</i>	12; this study

Crystallization, data collection, and processing

Purified SrtC was dialyzed against 50 mM Tris-HCl, 1 mM ethylenediaminetetraacetic acid and concentrated thereafter to 10 mg ml⁻¹. Two different crystallization conditions were established for SrtC. Orthorhombic P2₁2₁2₁ crystals appeared in sitting drop by vapor diffusion in 1.6 M magnesium sulfate and 0.1 M Mes (pH 6.5) at 20 °C (crystal form A) and were also obtained in hanging drops by vapor diffusion in 0.1 M Mes (pH 7.0) and 1.9 M ammonium sulfate at 20 °C (crystal form B). The two crystal structures from SrtC were here termed by their respective PDB codes, 3G66 (crystal form A) and 3G69 (crystal form B). Two microliters of a 10 mg/ml protein solution was mixed in 1 μl of the crystallization reservoir solutions for the hanging drops, and 0.1 μl of the same protein solution in 0.1 μl of the crystallization reservoir solutions for the sitting drops. Both crystal forms were

cryoprotected by adding 20% glycerol, followed by flash cooling in liquid nitrogen. Data collection was performed under cryogenic conditions (*T* = 100 K), at beamline ID14-1 in European Synchrotron Radiation Facility (Grenoble, France) (λ = 0.934 Å) using an ADSC Q210 CCD detector. The diffraction data were processed using MOSFLM.³⁶ Data collection statistics for the data sets as well as cell dimensions of the two orthorhombic crystals of SrtC are presented in Table 1.

Structure determination and refinement

Both crystal structures of SrtC were solved by molecular replacement using the PHASER program.³⁷ Matthew coefficient calculations suggested the presence of two molecules in the asymmetric units for both crystals.³⁸ The crystal structures of SrtC were determined using a polyalanine model of SrtB (58% identity and PDB code 2W1J). Successive cycles of manual model building were performed with COOT,³⁹ with the rigid-body refinement in the initial steps and with the restrained refinement in the later steps by using REFMAC.⁴⁰ Five percent of the reflections was set aside for monitoring the refinement by *R*_{free}.⁴¹ The loop (comprising residues 79–100) was built manually in both crystal structures of SrtC. Besides residues 255–267 at the C-terminal of SrtC, the loop residues 90–97 and 96–98 in chains A and B of SrtC (PDB code 3G66), respectively, could not be modeled due to the poor quality of the electron density. Similarly, we were not able to model the side chains of residues 89–98 and 90–93 in chains A and B of SrtC (PDB code 3G69), respectively. Two 2-(*N*-morpholino)ethanesulfonic acid molecules were added in observed excess densities within both SrtC crystal structures. The final rounds of refinement yielded *R*_{free} and *R*_{cryst} values of 23.5% and 18.7% for SrtC (PDB

Table 3. Primers used in the study, with restriction enzymes sites

Primer name	Sequence
SrtA-1	aggatgaggacttgattgaagaat
SrtA-2	tttgggcccctatgcttcacctctgttctgtt (ApaI site underlined)
SrtA-3	tttggatcctacaatcagtgaaatccatgatt (BamHI site underlined)
SrtA-4	tccgataaagttccgatgaaagt
SrtBT131Rfor	gtgattacggcacatagaggttgccaaacagct
SrtBT131Rrev	agctgttgccaaacctctatgtgcccgaatcac
Dam406	tctatgcctattccagaggaaatggat
Kan-Rev-Apa	tttgggcccctaaaacaattcatccagtaaaaat (ApaI site underlined)

code 3G66) and 25.8% and 19.2% for SrtC (PDB code 3G69), respectively. The two models conform well to the expected protein geometry and the rmsd of bond lengths from standard values for such resolutions (Table 1). The stereochemistry of the models was analyzed with PROCHECK.⁴² Refinement statistics are provided in Table 1. The secondary structure assignment was performed with DSSP.⁴³ The SSM algorithm was used to calculate rmsd.⁴⁴

Accession numbers

Coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers 3G66 and 3G69.

Acknowledgements

We gratefully acknowledge access to synchrotron radiation at beamline ID14-1 at the European Synchrotron Radiation Facility laboratory, Grenoble, France; at beamline X11 at the DESY laboratory, Hamburg, Germany; at beamline MX-41 at the BESSY laboratory, Berlin, Germany; and at beamline 711 at MAX laboratory, Lund University, Sweden. This study was supported by grants from the Swedish Research Council, the Swedish National Board of Health and Welfare, the Royal Academy of Sciences, Torsten and Ragnar Söderbergs Foundation, the Swedish Foundation for Strategic Research, and the German Research Council. The authors have no conflicting financial interests.

References

- Cutts, F. T., Zaman, S. M., Enwere, G., Jaffar, S., Levine, O. S., Okoko, J. B. *et al.* (2005). Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial. *Lancet*, **365**, 1139–1146.
- Sandgren, A., Sjöström, K., Olsson-Liljequist, B., Christensson, B., Samuelsson, A., Kronvall, G. & Henriques Normark, B. (2004). Effect of clonal and serotype-specific properties on the invasive capacity of *Streptococcus pneumoniae*. *J. Infect. Dis.* **189**, 785–796.
- Barocchi, M. A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A. *et al.* (2006). A pneumococcal pilus influences virulence and host inflammatory responses. *Proc. Natl Acad. Sci. USA*, **103**, 2857–2862.
- Hultgren, S. J., Jones, C. H. & Normark, S. (1996). Bacterial adhesins and their assembly. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., ed), pp. 2730–2756, American Society for Microbiology Press, Washington, DC.
- Budzík, J. M., Marraffini, L. A. & Schneewind, O. (2007). Assembly of pili on the surface of *Bacillus cereus* vegetative cells. *Mol. Microbiol.* **66**, 495–510.
- Mora, M., Bensì, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A. G. *et al.* (2005). Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. *Proc. Natl Acad. Sci. USA*, **102**, 15641–15646.
- Nallapareddy, S. R., Singh, K. V., Sillanpää, J., Garsin, D. A., Hook, M., Erlandsen, S. L. & Murray, B. E. (2006). Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. *J. Clin. Invest.* **116**, 2799–2807.
- Ton-That, H. & Schneewind, O. (2003). Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol. Microbiol.* **50**, 1429–1438.
- Yanagawa, R. & Honda, E. (1976). Presence of pili in species of human and animal parasites and pathogens of the genus *Corynebacterium*. *Infect. Immun.* **13**, 1293–1295.
- Butler, J. C., Hofmann, J., Cetron, M. S., Elliott, J. A., Facklam, R. R. & Breiman, R. F. (1996). The continued emergence of drug-resistant *Streptococcus pneumoniae* in the United States: an update from the Centers for Disease Control and Prevention's Pneumococcal Sentinel Surveillance System. *J. Infect. Dis.* **174**, 986–993.
- Sjöström, K., Blomberg, C., Fernebro, J., Dagerhamn, J., Morfeldt, E., Barocchi, M. A. *et al.* (2007). Clonal success of pilated penicillin nonsusceptible pneumococci. *Proc. Natl Acad. Sci. USA*, **104**, 12907–12912.
- Falkner, S., Nelson, A. L., Morfeldt, E., Jonas, K., Hulténby, K., Ries, J. *et al.* (2008). Sortase-mediated assembly and surface topology of adhesive pneumococcal pili. *Mol. Microbiol.* **70**, 595–607.
- Nelson, A. L., Ries, J., Bagnoli, F., Dahlberg, S., Falkner, S., Rounioja, S. *et al.* (2007). RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol. Microbiol.* **66**, 329–340.
- Mazmanian, S. K., Liu, G., Ton-That, H. & Schneewind, O. (1999). *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science*, **285**, 760–763.
- Schneewind, O., Model, P. & Fischetti, V. A. (1992). Sorting of protein A to the staphylococcal cell wall. *Cell*, **70**, 267–281.
- Fischetti, V. A., Pancholi, V. & Schneewind, O. (1990). Conservation of a hexapeptide sequence in the anchor region of surface proteins from Gram-positive cocci. *Mol. Microbiol.* **4**, 1603–1605.
- Mandlik, A., Swierczynski, A., Das, A. & Ton-That, H. (2008). Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol.* **16**, 33–40.
- Manzano, C., Contreras-Martel, C., El Mortaji, L., Izore, T., Fenel, D., Vernet, T. *et al.* (2008). Sortase-mediated pilus fiber biogenesis in *Streptococcus pneumoniae*. *Structure*, **16**, 1838–1848.
- Ton-That, H. & Schneewind, O. (2004). Assembly of pili in Gram-positive bacteria. *Trends Microbiol.* **12**, 228–234.
- Swaminathan, A., Mandlik, A., Swierczynski, A., Gaspar, A., Das, A. & Ton-That, H. (2007). House-keeping sortase facilitates the cell wall anchoring of pilus polymers in *Corynebacterium diphtheriae*. *Mol. Microbiol.* **66**, 961–974.
- LeMieux, J., Woody, S. & Camilli, A. (2008). Roles of the sortases of *Streptococcus pneumoniae* in assembly of the RlrA pilus. *J. Bacteriol.* **190**, 6002–6013.
- Neiers, F., Madhurantakam, C., Falkner, S., Normark, S., Henriques-Normark, B. & Achour, A. (2009). Cloning, expression, purification, crystallization and preliminary X-ray analysis of the pilus-associated sortase C from *Streptococcus pneumoniae*. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **65**, 55–58.

23. Bond, C. S. (2003). TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics*, **19**, 311–312.
24. Ilangovan, U., Ton-That, H., Iwahara, J., Schneewind, O. & Clubb, R. T. (2001). Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc. Natl Acad. Sci. USA*, **98**, 6056–6061.
25. Maresso, A. W., Wu, R., Kern, J. W., Zhang, R., Janik, D., Missiakas, D. M. *et al.* (2007). Activation of inhibitors by sortase triggers irreversible modification of the active site. *J. Biol. Chem.* **282**, 23129–23139.
26. Zhang, R., Wu, R., Joachimiak, G., Mazmanian, S. K., Missiakas, D. M., Gornicki, P. *et al.* (2004). Structures of sortase B from *Staphylococcus aureus* and *Bacillus anthracis* reveal catalytic amino acid triad in the active site. *Structure*, **12**, 1147–1156.
27. Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O. & Narayana, S. V. (2004). Crystal structures of *Staphylococcus aureus* sortase A and its substrate complex. *J. Biol. Chem.* **279**, 31383–31389.
28. Zong, Y., Mazmanian, S. K., Schneewind, O. & Narayana, S. V. (2004). The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the *Staphylococcus aureus* cell wall. *Structure*, **12**, 105–112.
29. Folch, B., Rooman, M. & Dehouck, Y. (2008). Thermostability of salt bridges *versus* hydrophobic interactions in proteins probed by statistical potentials. *J. Chem. Inf. Model.* **48**, 119–127.
30. Reid, K. C., Lindley, P. F. & Thornton, J. M. (1985). Sulphur-aromatic interactions in proteins. *FEBS Lett.* **190**, 209–213.
31. Viguera, A. R. & Serrano, L. (1995). Side-chain interactions between sulfur-containing amino acids and phenylalanine in α -helices. *Biochemistry*, **34**, 8771–8779.
32. Ton-That, H., Liu, G., Mazmanian, S. K., Faull, K. F. & Schneewind, O. (1999). Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc. Natl Acad. Sci. USA*, **96**, 12424–12429.
33. Matagne, A., Dubus, A., Galleni, M. & Frere, J. M. (1999). The β -lactamase cycle: a tale of selective pressure and bacterial ingenuity. *Nat. Prod. Rep.* **16**, 1–19.
34. Sung, C. K., Li, H., Claverys, J. P. & Morrison, D. A. (2001). An rpsL cassette, janus, for gene replacement through negative selection in *Streptococcus pneumoniae*. *Appl. Environ. Microbiol.* **67**, 5190–5196.
35. Lau, P. C., Sung, C. K., Lee, J. H., Morrison, D. A. & Cvitkovitch, D. G. (2002). PCR ligation mutagenesis in transformable streptococci: application and efficiency. *J. Microbiol. Methods*, **49**, 193–205.
36. Leslie, A. G. (2006). The integration of macromolecular diffraction data. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 48–57.
37. McCoy, A. J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **63**, 32–41.
38. Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497.
39. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
40. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F. & Murshudov, G. N. (2004). REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2184–2195.
41. Brunger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
42. Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049–1067.
43. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
44. Krissinel, E. & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2256–2268.

Résumé :

Streptococcus pneumoniae est un pathogène majeur chez l'homme, responsable d'otites sévères, de pneumonies, de bactériémies et de méningites. C'est une cause majeure de mortalité et de morbidité, essentiellement chez les enfants et les personnes âgées. Il a été récemment découvert que certaines souches virulentes de *S. pneumoniae* portent à leur surface des pili, considérés comme un important facteur de virulence. Cependant, contrairement aux bactéries à Gram-négatif, peu de choses sont connues sur la structure des pili des bactéries à Gram-positif. Les gènes requis pour la production des pili chez *S. pneumoniae* sont localisés sur un îlot de pathogénicité et codent pour 3 sortases (SrtC-1, SrtC-2, SrtC-3) et 3 protéines structurales (RrgB, qui forme le corps du pilus, RrgA et RrgC). Celles-ci contiennent un motif LPXTG, motif de reconnaissance des enzymes sortases.

Dans ce travail, nous apportons de nouvelles informations structurale, biochimique et microbiologique sur le mécanisme de formation du pilus par (1) une reconstitution du corps de la fibre *in vitro* après avoir identifié SrtC-1 comme étant la principale pilus-polymérase; (2) une étude de microscopie électronique montrant que les fibres produites *in vitro* mimaient structuralement les pili; (3) une étude *in vivo* confirmant les résultats obtenus *in vitro*; (4) la résolution de la structure par cristallographie aux rayons X de SrtC-1 et SrtC-3 qui révèle un site actif dont l'accès est contrôlé par un couvercle flexible, contrairement aux sortases non impliquées dans la biogenèse de pili. Ces observations suggèrent que la spécificité de substrat est dictée par une reconnaissance de surface couplée à une ouverture du couvercle. Dans un second temps, nous avons caractérisé la région du site actif de SrtC-1 : la triade catalytique mais aussi les deux résidus du couvercle qui s'ancrent dans le site actif. L'identification de résidus clés dans l'activité sortasique aussi bien que dans la stabilisation structurale a été ainsi mise en évidence.

Abstract :

Streptococcus pneumoniae, a pilated pathogen, is the causative agent of otitis, pneumonia, bacteremia, and meningitis, and is a major cause of community-acquired illnesses, especially in the very young and the elderly. Pili have been shown to play key roles in pneumococcal infection, and although the pilus biogenesis mechanism has been well studied in Gram-negative bacteria, it remains poorly understood in Gram-positive species. Pilus assembly in *S. pneumoniae* involves a membrane-associated macromolecular machinery encoded on a single pathogenicity islet, and includes the backbone fiber protein (RrgB), two minor pilins (RrgA and RrgC), and three dedicated sortases (SrtC-1, SrtC-2, SrtC-3). These pilin subunits contain the recognition LPXTG motif of sortases.

Here, we provide structural, biochemical, and microbiological insight into this novel pilus formation mechanism by (1) reconstituting the main backbone fiber of the pneumococcal pilus *in vitro* upon identification that SrtC-1 acts as the main pilus-polymerase and can form RrgB fibers; (2) employing electron microscopy in order to visualize fibers produced *in vitro* and showing that they structurally mimic the pneumococcal pilus backbone; (3) confirming this result *in vivo*; (4) solving the high resolution crystal structures of SrtC-1 and SrtC-3 which reveal active sites whose access is controlled by flexible lids, unlike in non-pilus sortases, suggesting that substrate specificity is dictated by surface recognition coupled to lid opening. In addition, we characterize the catalytic triad present in the sortase active site as well as lid residues, and identify key amino acids which are relevant for sortase activity as well as structural stabilization.